

# PATHOPHYSIOLOGICAL DOMAINS UNDERLYING THE METABOLIC SYNDROME: AN ALTERNATIVE FACTOR ANALYTIC STRATEGY

CAREL F.W. PEETERS\*, JAMES DZIURA, AND FLORYT VAN WESEL

**ABSTRACT.** Factor analysis (FA) has become part and parcel in metabolic syndrome (MBS) research. Both exploration- and confirmation-driven factor analyzes are rampant. However, factor analytic results on MBS differ widely. A situation that is at least in part attributable to misapplication of FA. Here, our purpose is (i) to review factor analytic efforts in the study of MBS with emphasis on misuse of the FA model and (ii) to propose an alternative factor analytic strategy.

*Methods:* The proposed factor analytic strategy consists of four steps and confronts weaknesses in application of the FA model. At its heart lies the explicit separation of dimensionality and pattern selection as well as the direct evaluation of competing inequality-constrained loading patterns. A high-profile MBS data set with anthropometric measurements on overweight children and adolescents is reanalyzed using this strategy.

*Results:* The reanalysis implied a more parsimonious constellation of pathophysiological domains underlying phenotypic expressions of MBS than the original analysis (and many other analyzes). The results emphasize correlated factors of impaired glucose metabolism and impaired lipid metabolism.

*Conclusions:* Pathophysiological domains underlying phenotypic expressions of MBS included in the analysis are driven by multiple interrelated metabolic impairments. These findings indirectly point to the possible existence of a multifactorial aetiology.

*Key words:* Factor analysis; Metabolic syndrome

*Abbreviations:* AACE = American Academy of Clinical Endocrinologists; BMI = body mass index; CFA = confirmatory factor analysis; CVD = cardiovascular disease; DBP = diastolic blood pressure; DM = diabetes mellitus; EFA = exploratory factor analysis; EGIR = European Group for study of Insulin Resistance; FA = factor analysis; G2 blood glucose level two hours after oral glucose intake; GB = blood glucose level at (fasting) baseline; HDL chol. = high density lipoprotein cholesterol; HOMA = homeostatic model assessment; IR = insulin resistance; MBS = metabolic syndrome; NCEP ATP III = National Cholesterol Education Program - Adult Treatment Panel III; PCA = principal components analysis; SBP = systolic blood pressure; trig. = triglycerides; UCFM = unrestricted confirmatory factor model; WHO = World Health Organization

---

\*Corresponding author.

## 1. INTRODUCTION

Certain risk factors for type 2 diabetes mellitus (DM) and atherosclerotic cardiovascular disease (CVD) have long been observed to cluster together in the individual [1, 2]. This clustering has rendered renewed interest with Reaven's contention [3, 4] that insulin resistance forms its basis. Today, the complex of interrelated risk factors of metabolic origin is known as the 'metabolic syndrome' (MBS) [5, 6]. It is considered to be a major threat to current and future public health, especially as MBS might result from maladaptive human metabolism in the face of food energy abundance in combination with a sedentary lifestyle [7, 8].

Recently, the MBS concept has been hotly debated [9–18]. A number of interrelated reasons are at the heart of the debate. The aetiology of the syndrome is largely unknown as to date it is unclear if a single pathogenetic process promotes the syndrome, or if multiple different pathogenetic processes need to concur in order for the syndrome to express itself. Notwithstanding the unknown aetiology, a number of expert groups, among which the World Health Organization (WHO) [19] and the National Cholesterol Education Program - Adult Treatment Panel III (NCEP ATPIII) [20], have published (slightly) different definitions of MBS intended for clinical diagnosis. However, clinical evidence on whether MBS is a better predictor of CVD and DM risk than its individual components, is equivocal [10, 21].

The stance on the MBS concept taken here is the following. The aetiology does not have to be known for the existence of a condition to be accepted [17], as is the case with, for example, type 2 DM. Current knowledge is however limited, such that MBS is not considered a clinical entity, but rather as defining a state of heightened risk for DM and CVD. The syndromic approach taken is epidemiological rather than clinical. In the sense that MBS is deemed to provide a conceptual framework for the clustering of metabolic risk factors. An important step in furthering epidemiologic understanding of the syndrome is then an assessment of the pathophysiological constellation of what are deemed to be phenotypic expressions of MBS. A constellation we believe to be driven, not by insulin resistance [3, 4], but by multiple interrelated metabolic impairments [22].

Factor analysis (FA) has become an oft-used tool for evaluations of phenotypic domains underlying MBS [23]. FA is a multivariate technique that may reveal a pattern of reduced dimensionality among a larger set of intercorrelated variables [24]. FA, however, is often poorly understood while being routinely executed [25], leading to diffuse findings and possibly confusing efforts in understanding MBS. This paper aims to review factor analytic efforts in the study of MBS, with emphasis on misuse of FA. An alternative factor analytic strategy is proposed that confronts weaknesses in the application of FA. A high-profile MBS data set with anthropometric measurements on overweight and obese children and adolescents is reanalyzed using the alternative strategy. The findings may give renewed cachet to both FA and its connection to MBS research.

## 2. ASSESSING FACTOR ANALYTIC EFFORTS IN MBS RESEARCH

**2.1. The Factor Analytic Model.** FA has come to be heavily utilized in the MBS research community since its seminal usage by Edwards, et al. [26]. An important question thrusting the FA efforts in MBS research is if a unifying physiology dominated by insulin resistance underlies the clustering of metabolic risk variables, or if there are multiple underlying physiologic phenotypes.

The common factor analytic model assumes that a random  $p$ -dimensional vector of observed variables can be grouped by their covariances or correlations into a lower-dimensional linear combination of latent variables:

$$(1) \quad \underset{(p \times 1)}{\mathbf{z}_i} = \underset{(p \times 1)}{\boldsymbol{\mu}} + \underset{(p \times m)}{\boldsymbol{\Lambda}} \cdot \underset{(m \times 1)}{\boldsymbol{\xi}_i} + \underset{(p \times 1)}{\boldsymbol{\epsilon}_i}.$$

In (1)  $\mathbf{z}_i$  denotes the (possibly standardized) observed variable of dimension  $p$  for person  $i$ ,  $\boldsymbol{\mu}$  denotes the intercept,  $\boldsymbol{\epsilon}_i$  denotes the error measurements for person  $i$ , and  $\boldsymbol{\Lambda}$  is a  $(p \times m)$ -dimensional matrix of factor loadings in which each element  $\lambda_{jk}$  is the loading of the  $j$ th variable on the  $k$ th factor,  $j = 1, \dots, p$ ,  $k = 1, \dots, m$ . Then  $\boldsymbol{\xi}_i$  represents a latent variable of dimension  $m$ , with  $m < p$ , whose elements are referred to as common factors. In effect, FA represents a method of identifying or specifying latent factors that account for the (co)variances among observed variables by partitioning observed variance into common variance (attributable to the underlying latent common factors) and unique variance (among which are error components) [24]. In the standard model the random variables  $\mathbf{z}_i$ ,  $\boldsymbol{\xi}_i$ , and  $\boldsymbol{\epsilon}_i$  are assumed to have Gaussian distributions, although the development of robust estimation procedures (e.g., [27, 28]) have somewhat softened the necessity of this assumption. (See Appendix ‘Basics of Factor Analysis’ for a more detailed overview of the FA model). The model can have both explorative and confirmative thrusts. Before assessing these thrusts some general remarks on FA are made.

A first general comment on the utilization of FA by the MBS research community is that basic model assumptions are seldom assessed. The model boasts several implicit assumptions such as a nonsingular sample covariance matrix and a reasonable proportion of variance among the observed variables being common variance. The appropriateness of these assumptions is easily [29] but rarely assessed. Moreover, the explicit distributional assumptions imply usage of observed variables of continuous metric and disqualifies the common FA model for binary and categorical observed data. Many MBS studies, however, employ standard FA on variables of non-continuous metric. In such situations extensions of the standard FA model are needed [30, 31].

A second general comment concerns interpretational overextension. The FA model cannot determine existence of MBS nor assess clinical importance of MBS as a concept [25]. What FA *can* do is, through the latent factors (when adhering to a realist ontology [32]), give indications of pathophysiological domains that underlie phenotypic expressions of MBS.

**2.2. Comments on Exploratory Efforts.** In exploratory FA (EFA) both  $m$  and the meaning of latent factors are unknown. In the exploratory sense, FA is a theory-generating technique used for the identification of meaningful latent factors. Most MBS studies utilize FA in the exploratory sense [33–43]. Many deployments of EFA are however suboptimal.

EFA is often confused with principal components analysis (PCA). PCA is a data-reductional technique, seeking to identify components. It resembles the model in (1) without inclusion of error measurements, leading the components to differ conceptually and mathematically from latent factors in EFA [44]. Components are weighted linear combinations of observed variables seeking to efficiently explain observed variance in the data [45], leaving the explanation of observed covariance

secondary [44]. Many MBS studies employing FA, however, seek to obtain an explanation of the observables' covariation through a small number of explanatory factors. Also, a phenotype is just an expression of genotype or pathophysiology, indicating the necessity of including measurement error. As such, employing common EFA would be more appropriate.

An important decision in EFA is determining the number of factors to retain. Many methods are heuristical, relying on subjective judgments or arbitrary cut-off values. The Guttman-Kaiser rule [46, 47] is the most popular rule of thumb. It states that one should retain at most those factors associated with eigenvalues whose magnitude exceeds the average eigenvalue (the average eigenvalue being 1 when using standardized data). This criterion, as well as many other heuristical factor retention criteria, are prone to under- (retaining too few factors) and overfactoring (retaining too many factors) [48, 49].

Given the above it is disconcerting that factor analytic efforts in MBS research are usually based on what has been termed 'Little jiffy' [29]: Employment of PCA, retainment of components based on the Guttman-Kaiser rule followed by a Varimax rotation [50], and the subsequent interpretation of rotated components as if they were common factors. Such mechanical use of EFA stunts learning and interpretation [51], and is (in part) responsible for the widely differing results obtained with EFA in MBS research. See the online supplement to this paper for an overview of the disparate results obtained with EFA as well as confirmatory FA (next subsection).

**2.3. Comments on Confirmatory Efforts.** Confirmatory FA (CFA) is a theory-testing technique. An *a priori* factor structure is assumed, with given  $m$ , with a pre-specified loadings matrix in which exclusion constraints indicate which variables are indicators of which latent factor(s), and with possibly correlated factors and error variances. The model or models stated then remain to be tested. CFA studies are gaining interest in MBS research [52–58]. Standard CFA however, can also be misapplied.

The evaluation of model fit in CFA is essentially the evaluation of a diffuse hypothesis as it is unclear in case of misspecification if the pattern of loadings or the factor dimensionality is to blame [59]. Moreover, specifying a pattern of factor loadings through exclusion restrictions implies a loss of information in the sense that more exclusion restrictions are applied than is usually necessary for identification of the FA model. Additionally, exclusion restrictions may amount to errors of omission, may make the unrealistic assumption that items are factorially pure (in the population), and may induce bias in estimates of the free parameters [60, 61]. These issues are intricately connected to the well-known and widespread situation of exploratively obtained factor structures not being confirmed by CFA [60].

Some recent CFA studies in MBS research claim to provide evidence that a single latent factor underlies MBS [53, 55, 58]. These studies include four observed variables, some of which are functions of several variables usually employed as separate phenotypic items. This practice (called parceling) is justified by the claim that utilization of multiple measures for what is believed to be a trait will lead to a model with multiple factors, thus clouding efforts to establish a single latent factor underlying MBS. However, the inclusion of multiple sets of correlated measures does not irrevocably lead to a model with more than one latent factor,

unless some measures would identify a doublet factor (see Section ‘Step 1: Dimensionality Selection’ below). Also, the mentioned CFA efforts may actually provide evidence for a hierarchical latent factor [52] rather than a single pathophysiological domain underlying MBS, as the usage of functions of variables implies a (partial) pre-compression of the data. Parceling changes the nature of the data, may mask model misspecification and thus may inflate goodness-of-fit (indices) [62] (cf. [63]). Indeed, “it is possible to argue (inappropriately) that a single factor solution provides a good fit to the data under apparently benign parceling strategies” [62]. Moreover, while a two-factor model can be modeled on four variables, such a model would not be meaningful given that the observables were constructed to represent a single factor; implying that the one-factor model is the only model to be meaningfully fitted on the four observed variables. The possibility to assess if there are multiple (related) pathophysiological domains underlying phenotypic expressions of MBS is then denied. A meaningful scientific method, however, allows for multiple competing theories to be tested [64], as single hypotheses suffer from confirmation bias.

### 3. AN ALTERNATIVE FACTOR ANALYTIC STRATEGY

An alternative confirmatory factor analytic strategy is proposed that aims to confront the weaknesses in the application of FA. In a sense the strategy seeks to bridge the EFA – CFA divide so as to increase the inferential power of a factor analysis. This strategy is embedded within the Bayesian model selection approach, whose analytical advantages have been well documented [65]. The main Bayesian model selection criterion is the Bayes factor. This quantity incorporates model fit as well as model complexity and expresses “the evidence provided by the data in favor of one scientific theory, represented by a statistical model, as opposed to another” [66]. It can be used to compare any two models. Using Bayes factors, one can compute posterior model probabilities, which, assuming a uniform prior on the model space, are normalized Bayes factors. These quantities express model (un)certainty, in the sense that posterior model probabilities can be interpreted as the relative amounts of support in the data for the models under consideration. See Appendix ‘A Primer on Bayesian Statistics’ for more information on the Bayes factor and posterior model probabilities.

The strategy proposed consists of the following steps:

- Step 1:** Determine formally the (intrinsic) number of common factors based on a weighing of model fit and model complexity. For example along the lines specified in Peeters [67], who uses a Bayesian EFA model for selecting the optimal dimension  $m$ ;
- Step 2:** Whence settled on latent factor dimensionality  $m$ , specify an unrestricted confirmatory factor model (UCFM). An UCFM is a FA model that corresponds to EFA in the sense that only minimal restrictions are placed on the factor loadings matrix and the factor covariance matrix for achieving global rotational uniqueness of the factor solution. However, the restrictions are to be chosen such that they convey preconceived theoretical meaning and thus render unnecessary post-hoc rotation of the solution for interpretation purposes. Peeters [68] gives minimal conditions for specifying a UCFM;

**Step 3:** Formulate, using the UCFM obtained in Step 2 as a base model, competing inequality constrained factor structures making use of inequality constraints on and between the free parameters in the loadings matrix. Substantive theory is then not represented by exclusion restrictions to express a pre-specified factor loading pattern, but by the imposition of inequality constraints;

**Step 4:** Compute the posterior model probability for each constrained model under consideration and determine the constrained model most supported by the data.

The strategy explicitly encourages the formulation of competing inequality constrained theories for (statistical) scrutiny. When used in full, the thrust of the sequence is confirmatory, with the explicit separation of dimensionality and pattern selection in order to avoid embarking on diffuse hypotheses.

#### 4. REANALYSIS OF DATA BY WEISS ET AL. [69]

**4.1. Data.** The data have been described elsewhere [69]. It considers a multiethnic, multiracial cohort of 464 nondiabetic obese and overweight children and adolescents. The data contain measurements on the body mass index (BMI), blood glucose level at (fasting) baseline (GB) and two hours after (G2) oral glucose intake (both in mg/dl), fasting levels of triglycerides (trig.; mg/dl) and high-density lipoprotein (HDL) cholesterol (mg/dl), systolic and diastolic blood pressure (SBP, DBP; both in mm Hg), and insulin resistance (IR). IR was measured through homeostatic model assessment (HOMA). For more information on the data, see Weiss et al. [69].

Consideration of pediatric samples is important as current prevalence of MBS among youngsters may give indications of the future burden of DM and CVD. The measurements are in line with the American Academy of Clinical Endocrinologists (AACE) position on MBS [70], which emphasizes the epidemiologic pathophysiological perspective. The inclusion of IR makes it possible to test theories regarding the importance of IR in the MBS construct.

The little jiffy approach was the original factor analytic strategy for analyzing the data [69]. Here, the alternative strategy will be utilized for reanalysis. As in Weiss et al. [69], the natural logarithm was taken of the glucose, insulin resistance and triglycerides measurements to abide the normality assumption. The data were standardized such that a case of modeling the correlation matrix is considered. The sample correlation matrix is nonsingular and the Kaiser-Meyer-Olkin test [29] indicates that a reasonable proportion of variance among the variables might be common variance.

**4.2. Step 1: Dimensionality Selection.** Posterior model probabilities are computed for each model allowed by the condition  $(p-m)^2 - p - m \geq 0$ . This inequality simply states that the number of nonredundant elements in the sample covariance matrix must be greater than or equal to the number of freely estimable parameters in the model, which places an upper bound on  $m$ . The data have eight measured variables ( $p = 8$ ), giving that the maximum number of factors that can be extracted equals  $m = 4$ . The computation strategy couples the candidate estimator method for computing Bayes factors [71] with the use of training samples [72]. This practice allows one to obtain determinate Bayes factors and subsequent posterior model probabilities using standard diffuse conjugate or noninformative priors [73]. The

computation strategy is embedded within a search strategy too weed out models suffering from rank deficiency in  $\mathbf{\Lambda}$  as this is a direct indicator for overfactoring (see (6) in Appendix ‘Basics of Factor Analysis’). This search excludes the  $m = 4$  model. The following posterior model probabilities are obtained when assuming that each model is equally likely *a priori*:  $P(m = 1|\mathbf{Z}) = 0$ ;  $P(m = 2|\mathbf{Z}) = 1$ ; and  $P(m = 3|\mathbf{Z}) = 0$ . The data thus support the two-factor model.

These results differ from Weiss et al. [69] and several other factor analytic efforts in which a three-factor (or higher) solution was found. A first reason for the retention of more latent factors in these studies is the tendency of heuristic factor-selection rules to overfactor [48, 49]. More formal selection procedures, such as the likelihood ratio test in maximum likelihood EFA and the assessment of information criteria, do not escape this tendency [74, 75].

Another reason for the higher factor solution in other studies might be the existence of doublet factors. Doublet factors are factors that arise as the result of common variance due to correlation between just two variables [24]. Doublet factors are considered to be conceptually weak factors and the assignment of an independent latent construct to a doublet factor is contentious. SBP and DBP are variables that, within the battery of measurements on phenotypic expressions of MBS, usually correlate only with each other resulting in a doublet factor to be extracted (usually termed ‘hypertension’). The strategy employed here sees past the doublet factor and indicates the more parsimonious model with two factors as optimal.

**4.3. Step 2: Base Model Formulation.** A UCFM for  $m = 2$  will be formulated for confirmatory efforts. Abiding conditions given by Peeters [68], the following minimal restrictions on  $\mathbf{\Lambda}$  are chosen for global rotational uniqueness:

$$\mathbf{\Lambda}_0 = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} = 0 & \lambda_{32} > 0 \\ \lambda_{41} & \lambda_{42} \\ \lambda_{51} > 0 & \lambda_{52} = 0 \\ \lambda_{61} & \lambda_{62} \\ \lambda_{71} & \lambda_{72} \\ \lambda_{81} & \lambda_{82} \end{bmatrix} \begin{array}{l} \text{BMI} \\ \log_e\{\text{trig.}\} \\ \text{HDL chol.} \\ \log_e\{\text{IR}\} \\ \log_e\{\text{GB}\} \\ \log_e\{\text{G2}\} \\ \text{SBP} \\ \text{DBP} \end{array}.$$

The exclusion restrictions  $\{\lambda_{31}, \lambda_{52}\} = 0$  identify the model up to polarity reversals in the columns. The polarity truncations  $\{\lambda_{32}, \lambda_{51}\} > 0$  then ensure global rotational uniqueness of the model. These constraints are chosen for the following reasons. First, from prior knowledge and previous analyzes a two-factor solution is deemed to consist of a glucose and a lipid factor. HDL chol. is then believed to have a large loading on a lipid (second) factor while having a small loading on the first factor, and  $\log_e\{\text{GB}\}$  is believed to have a large loading on a glucose (first) factor while having a small loading on the second factor. These variables then serve as an indicator of the respective factors. It is thus reasonable to specify  $\{\lambda_{31}, \lambda_{52}\} = 0$  and  $\{\lambda_{32}, \lambda_{51}\} > 0$ . Second, the chosen minimal restrictions comply with all competing inequality constrained formulations of factor structure to be assessed.

**4.4. Step 3: Formulating Competing Constrained Factor Structures.** Factor structure for confirmatory efforts is not represented using exclusion restrictions

but by imposing inequality constraints on and between the parameters left free in the UCFM. The following inequality constrained competing factor structures are formulated:

$$\mathbf{A}_1 = \left[ \begin{array}{cc} \lambda_{11} & > & |\lambda_{12}| \\ |\lambda_{21}| & < & -\lambda_{22} \\ \lambda_{31} = 0 & & \lambda_{32} > 0 \\ \lambda_{41} & > & |\lambda_{42}| \\ \lambda_{51} > 0 & & \lambda_{52} = 0 \\ \lambda_{61} & > & |\lambda_{62}| \\ \lambda_{71} & > & |\lambda_{72}| \\ |\lambda_{81}| & < & -\lambda_{82} \end{array} \right] \begin{array}{l} \text{BMI} \\ \log_e\{\text{trig.}\} \\ \text{HDL chol.} \\ \log_e\{\text{IR}\} \\ \log_e\{\text{GB}\} \\ \log_e\{\text{G2}\} \\ \text{SBP} \\ \text{DBP} \end{array},$$

$$\mathbf{A}_2 = \left[ \begin{array}{cc} |\lambda_{11}| & < & -\lambda_{12} \\ |\lambda_{21}| & < & -\lambda_{22} \\ \lambda_{31} = 0 & & \lambda_{32} > 0 \\ \lambda_{41} > .4 & & \lambda_{42} < -.4 \\ \lambda_{51} > 0 & & \lambda_{52} = 0 \\ \lambda_{61} & > & |\lambda_{62}| \\ \lambda_{71} & < & -\lambda_{72} \\ \lambda_{81} & < & -\lambda_{82} \end{array} \right] \begin{array}{l} \text{BMI} \\ \log_e\{\text{trig.}\} \\ \text{HDL chol.} \\ \log_e\{\text{IR}\} \\ \log_e\{\text{GB}\} \\ \log_e\{\text{G2}\} \\ \text{SBP} \\ \text{DBP} \end{array},$$

$$\mathbf{A}_3 = \left[ \begin{array}{cc} \lambda_{11} & > & |\lambda_{12}| \\ |\lambda_{21}| & < & -\lambda_{22} \\ \lambda_{31} = 0 & & \lambda_{32} > 0 \\ \lambda_{41} & > & |\lambda_{42}| \\ \lambda_{51} > 0 & & \lambda_{52} = 0 \\ \lambda_{61} & > & |\lambda_{62}| \\ |\lambda_{71}| < .3 & & |\lambda_{72}| < .3 \\ |\lambda_{81}| < .3 & & |\lambda_{82}| < .3 \end{array} \right] \begin{array}{l} \text{BMI} \\ \log_e\{\text{trig.}\} \\ \text{HDL chol.} \\ \log_e\{\text{IR}\} \\ \log_e\{\text{GB}\} \\ \log_e\{\text{G2}\} \\ \text{SBP} \\ \text{DBP} \end{array}.$$

A formulation like  $\lambda_{71} < -\lambda_{72}$  states that the negative of  $\lambda_{72}$  is believed to be larger than  $\lambda_{71}$ . Note that this is a much more informative formulation than the more usual strategy of setting  $\lambda_{71} = 0$  and letting  $\lambda_{72}$  be free to be estimated in order to express the belief that SBP is an indicator for the second latent factor rather than the first one. In the same respect, a formulation like

$$\lambda_{61} > |\lambda_{62}| \Rightarrow \begin{cases} \lambda_{61} - \lambda_{62} > 0 \\ \lambda_{61} + \lambda_{62} > 0 \end{cases},$$

indicates the belief that  $\lambda_{61}$  is larger than  $\lambda_{62}$ , irrespective of the latter's sign. A statement like

$$|\lambda_{82}| < .3 \Rightarrow -.3 < \lambda_{82} < .3,$$

indicates the belief that  $\lambda_{82}$  takes a value in the interval  $[-.3, .3]$ .

In all models insulin resistance, blood glucose level at baseline and two hours after glucose intake are mainly related to the glucose factor, while fasting levels of triglycerides and HDL cholesterol form the base of the lipid factor. Note that for the lipid factor polarity fixation is brought about by demanding  $\lambda_{32} > 0$ . HDL cholesterol is generally regarded as 'good' cholesterol, meaning that the choice  $\lambda_{32} > 0$  amounts to modeling a factor denoting unimpaired lipid metabolism. Hence



formulations like  $|\lambda_{21}| < -\lambda_{22}$ , as under given polarity truncation the triglycerides item is believed to be strongly negatively related to a lipid factor.

Model 1 adds detail to the base model and the basic factors by stating that BMI and systolic blood pressure are linked to the glucose factor ( $\lambda_{11} > |\lambda_{12}|$ ,  $\lambda_{71} > |\lambda_{72}|$ ), while diastolic blood pressure is believed to be linked to the lipid factor ( $|\lambda_{81}| < -\lambda_{82}$ ). Model 2 states the hypothesis that BMI is an indicator for the lipid rather than the glucose factor ( $|\lambda_{11}| < -\lambda_{12}$ ). Also, in this model both systolic and diastolic blood pressure are related to the lipid rather than the glucose factor, with the additional belief that both blood pressure measures will load positively on the latter ( $\lambda_{71} < -\lambda_{72}$ ,  $\lambda_{81} < -\lambda_{82}$ ). Moreover, the second model states that insulin resistance may be the measure tying MBS together. In a multifactor model this would imply that the insulin resistance measure achieves a large or dominant loading on both the glucose and lipid factor ( $\lambda_{41} > .4$ ,  $\lambda_{42} < -.4$ ). Model 3 resembles the first model, but states that the association of systolic and diastolic blood pressure with the factors is rather loose.

**4.5. Step 4: Constrained-Model Selection and Interpretation.** Bayes factors for models under inequality constraints are easily computed [76]. Again, standard (diffuse) conjugate and noninformative priors are utilized. Assuming a uniform prior on the model space the following posterior model probabilities are obtained for the constrained two-factor models under consideration:  $P(M_1|\mathbf{Z}) = .0004$ ;  $P(M_2|\mathbf{Z}) = 0$ ; and  $P(M_3|\mathbf{Z}) = .9996$ . Conditioned on the data, the third model receives almost all support.

The third constrained model connects trig. and HDL chol. to a lipid metabolism factor and IR, GB, and G2 to a glucose metabolism factor. Moreover, the model states that BMI is related to the glucose metabolism factor rather than the lipid metabolism factor. The finding that blood pressure is not an independent pathophysiological factor is consistent with epidemiologic evidence that insulin resistance and lipid metabolism play a role in the pathogenesis of hypertension rather than hypertension being a physiologic phenotype [22].

TABLE 1. Posterior Means and 95% Credible Intervals for  $\Lambda_0$

Factor 1			Factor 2			Item
Parameter	Mean	95% CI	Parameter	Mean	95% CI	
$\lambda_{11}$	.324	[.207, .440]	$\lambda_{12}$	-.068	[-.191, .055]	BMI
$\lambda_{21}$	-.006	[-.303, .215]	$\lambda_{22}$	-.653	[-.956, -.379]	$\log_e\{\text{trig.}\}$
$\lambda_{31}$	-	-	$\lambda_{32}$	.706	[.442, .940]	HDL chol.
$\lambda_{41}$	.767	[.613, .921]	$\lambda_{42}$	-.179	[-.343, -.022]	$\log_e\{\text{IR}\}$
$\lambda_{51}$	.470	[.360, .585]	$\lambda_{52}$	-	-	$\log_e\{\text{GB}\}$
$\lambda_{61}$	.355	[.205, .492]	$\lambda_{62}$	-.124	[-.289, .036]	$\log_e\{\text{G2}\}$
$\lambda_{71}$	.274	[.136, .416]	$\lambda_{72}$	.029	[-.118, .171]	SBP
$\lambda_{81}$	.202	[.069, .347]	$\lambda_{82}$	.139	[-.017, .292]	DBP

The estimates of the UCFM given in Table 1 also lend support to model 3. (Confer Appendix ‘Results Factor Analytic Approach Weiss et al. [69]’ to see, in contrast, the parameter estimates obtained with the little jiffy approach of the original study [69]). The table contains posterior means and credible intervals. A

credible interval is a posterior probability interval. For example, the posterior probability that  $\lambda_{11}$  lies in the interval  $[-.207, .440]$  is .95. (See Appendix ‘Reproduced Correlation Structure’ for an indication of the success of the two-factor UCFM in retrieving the sample correlation matrix).

The estimates indicate that the hypertension variables seem to be relatively weak in their association with the respective factors. The credibility intervals indicate that these variables are mostly related to the glucose metabolism factor, which would be in line with the hypothesis that hypertension is related to insulin resistance and impaired glucose metabolism [3, 4]. Regarding the contention that insulin resistance is the basis for MBS: IR is related to both the glucose and lipid factors. However, the posterior mean of the loading tying IR to the lipid metabolism factor is relatively small and the upper bound of its credibility interval approaches zero. The second inequality constrained model is thus rightly not supported by the data.

The two latent factors are appreciably correlated with a posterior mean of  $-.277$  and a 95% credible interval of  $[-.417, -.137]$ . Note that, as stated in Section ‘Step 3: Formulating Competing Constrained Factor Structures’, the second factor models unimpaired lipid metabolism. Thus, the selected model indicates that, given the data and measurements, impaired glucose metabolism and impaired lipid metabolism are positively related pathophysiological domains. Moreover, the two factors connect to two main hypotheses regarding syndrome aetiology, stating that the risk-factor associations are due to abnormality of the insulin/glucose metabolism and/or abnormality of the lipid metabolism [22].

## 5. DISCUSSION

Applications of FA in MBS research were evaluated. Lacunae in both EFA and CFA were discussed. It is argued that the mechanical use of EFA and misunderstandings of CFA are, at least in part, responsible for the widely differing results obtained with FA on data with phenotypic expressions of MBS.

An alternative factor analytic strategy is proposed. The strategy consists of four steps and aims to confront the weaknesses in application of the FA model, by: (i) Formally assessing optimal choice of factor dimensionality; (ii) Canceling the need for post-hoc rotation of the factor solution; (iii) Allowing to express a confirmatory factor structure through informative inequality constraints rather than through rigid exclusion restrictions; (iv) Encouraging the formulation of competing inequality constrained theory-based expressions of factor structure, in order to avoid confirmation bias.

The alternative strategy was utilized in reanalyzing a high-profile data set on which factor analyses were previously employed. The data consider eight variables as phenotypic expressions of MBS in a cohort of nondiabetic overweight and obese children and adolescents [69]. The reanalysis based on the alternative strategy implied a more parsimonious constellation of pathophysiological domains underlying phenotypic expressions of MBS than the original analysis (and many other analyses). The selected two-factor solution stresses correlated factors of impaired glucose metabolism and impaired lipid metabolism. This solution does not assign hypertension a separate factor which is consistent with epidemiologic evidence that insulin resistance and lipid metabolism play a role in the pathogenesis of hypertension rather than hypertension being a physiologic phenotype [22]. Moreover, there

is no strong evidence of insulin resistance being dominant in both the glucose and lipid domains.

Several limitations of this study should be considered. First, the data consider a multiethnic cohort while MBS may express itself differently across ethnic groups and gender. Note, however, that the specification of a factor model through inequality constraints would also be helpful in assessing measurement (factorial) invariance across groups. Second, MBS may develop with age and with the advent of DM and CVD, implying that the data might represent a snap-shot of phenotypic expressions related to MBS. Third, the proposed factor analytic strategy is more involved than routine uses of EFA and regular CFA in the sense that it requires more computation time and puts higher cognitive demands on the researcher. Nevertheless, these drawbacks are felt to be outweighed by the advantages of the strategy and the proposed steps are deemed to form a viable analytic alternative for other studies seeking to use FA.

The findings suggest that there are two correlated pathophysiological domains underlying the phenotypic expressions of MBS included in the analysis. These domains are characterized by impaired glucose metabolism and impaired lipid metabolism, respectively. These findings indirectly point to the possibility that several different pathogenic processes need to coincide in order to be able to identify a MBS construct. It might be timely to postulate the possible existence of a multifactorial aetiology.

#### ACKNOWLEDGEMENTS

This work was supported by grant NWO-VICI-453-05-002 of the Netherlands Organization for Scientific Research (NWO). It was written while the author was a Ph.D. candidate at the Department of Methodology and Statistics, Utrecht University, Utrecht, the Netherlands and is part of the authors' Ph.D. thesis. This version is a postprint of: Peeters, C.F.W., Dziura, J., & van Wesel, F. (2014). Pathophysiological Domains Underlying the Metabolic Syndrome: An Alternative Factor Analytic Strategy. *Annals of Epidemiology*, 24: 762–770. This postprint is released under a Creative Commons BY-NC-ND license.

#### APPENDIX: BASICS OF FACTOR ANALYSIS

The unrestricted factor model is considered. Let  $\mathbf{Z}^T \equiv [\mathbf{z}_1, \dots, \mathbf{z}_n]$  define (standardized)  $p$ -variate observation vectors on  $i = 1, \dots, n$  subjects, such that  $\mathbf{z}_i^T \equiv [z_{i1}, \dots, z_{ip}] \in \mathbb{R}^p$  denotes a realization of the random vector  $Z_i^T \equiv [Z_{i1}, \dots, Z_{ip}] \in \mathbb{R}^p$ . Also, let  $\Xi^T \equiv [\xi_1, \dots, \xi_n]$  define  $m$ -variate vectors of latent factor scores on  $n$  subjects with  $\xi_i^T \equiv [\xi_{i1}, \dots, \xi_{im}] \in \mathbb{R}^m$ .

The model (1) maintains the following assumptions: (i)  $\mathbf{z}_i \perp\!\!\!\perp \mathbf{z}_{i'}, \forall i \neq i'$ ; (ii)  $\text{rank}(\Lambda) = m$ ; (iii)  $\epsilon_i \sim \mathcal{N}_p(\mathbf{0}, \Psi)$ , with  $\Psi \equiv \text{diag}(\psi_{11}, \dots, \psi_{pp})$ , and  $\psi_{jj} > 0$ ; (iv)  $\xi_i \sim \mathcal{N}_m(\mathbf{0}, \Phi)$ ; and (v)  $\xi_i \perp\!\!\!\perp \epsilon_{i'}, \forall i, i'$ . The likelihood for the observations conditional on the realization of  $\Xi$  can then be expressed as:

$$\begin{aligned}
 L(\mu, \Lambda, \Xi, \Psi, \Phi; \mathbf{Z}) &= \prod_{i=1}^n f(\mathbf{z}_i | \mu, \Lambda, \xi_i, \Psi, \Phi) \\
 (2) \qquad \qquad \qquad &= \prod_{i=1}^n (2\pi)^{-\frac{p}{2}} |\Psi|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \epsilon_i^T \Psi^{-1} \epsilon_i \right\},
 \end{aligned}$$

where  $\epsilon_i = \mathbf{z}_i - \boldsymbol{\mu} - \boldsymbol{\Lambda}\boldsymbol{\xi}_i$ . Marginalizing over  $\boldsymbol{\xi}_i$  the likelihood of the observed data can be obtained:

$$\begin{aligned}
 L(\boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\Psi}, \boldsymbol{\Phi}; \mathbf{Z}) &= \prod_{i=1}^n \int f(\mathbf{z}_i | \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\xi}_i, \boldsymbol{\Psi}, \boldsymbol{\Phi}) g(\boldsymbol{\xi}_i | \boldsymbol{\Phi}) \partial \boldsymbol{\xi}_i \\
 (3) \quad &= \prod_{i=1}^n (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi}|^{-\frac{1}{2}} \\
 &\quad \times \exp \left\{ -\frac{1}{2} (\mathbf{z}_i - \boldsymbol{\mu})^T [\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi}]^{-1} (\mathbf{z}_i - \boldsymbol{\mu}) \right\},
 \end{aligned}$$

giving that the factor decomposition constrains the covariance structure of the  $\mathbf{z}_i$  to

$$(4) \quad \boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi}.$$

Then, for existence (vi), generally  $(p - m)^2 - p - m \geq 0$ , simply stating that the number of nonredundant elements in the sample correlation matrix  $\mathbf{S}$  must be greater than or equal to the number of freely estimable parameters in  $\boldsymbol{\Sigma}$ , which places an upper bound on  $m$ .

Now,  $\boldsymbol{\Phi} \in \mathbb{R}^{m \times m}$  denotes the factor covariance matrix, giving that (4) represents an oblique model in which the latents may share covariation. Note that, for positive definite  $\boldsymbol{\Phi}$ , we may always find  $\mathbf{V} \in \mathbb{R}^{m \times m}$  such that  $\boldsymbol{\Phi} = \mathbf{V}\mathbf{V}^T$ , and

$$(5) \quad \boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi} = (\boldsymbol{\Lambda}\mathbf{V})[\mathbf{V}^{-1}\boldsymbol{\Phi}(\mathbf{V}^{-1})^T](\boldsymbol{\Lambda}\mathbf{V})^T + \boldsymbol{\Psi} = (\boldsymbol{\Lambda}\mathbf{V})(\boldsymbol{\Lambda}\mathbf{V})^T + \boldsymbol{\Psi}.$$

Equation (5) implies that any oblique representation has equivalent orthogonal representations. The orthogonal representation makes the following statements on identification less involved.

It is well known that for given  $\boldsymbol{\Lambda}$  and  $\boldsymbol{\Psi}$ , the former is defined uniquely only up to rotation. Correspondingly the FA literature has focussed mainly on identification of  $\boldsymbol{\Psi}$ . The main result of which is that if assumption (vi) holds,  $\boldsymbol{\Psi}$  is almost surely identified [77]. This result is however contingent upon the rank of  $\boldsymbol{\Lambda}$ . The implications of a failure to abide model assumption (ii) were explored by Anderson and Rubin [78] and Geweke and Singleton [79]. Suppose that  $\text{rank}(\boldsymbol{\Lambda}) = r < m$ . Then there exists a matrix  $\mathbf{Q} \in \mathbb{R}^{m \times (m-r)}$  for which  $(\boldsymbol{\Lambda}\mathbf{V})\mathbf{Q} = \mathbf{0}$  and  $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_{m-r}$ , such that for any  $\mathbf{M} \in \mathbb{R}^{p \times (m-r)}$  with mutually orthogonal rows

$$(6) \quad \boldsymbol{\Sigma} = (\boldsymbol{\Lambda}\mathbf{V})(\boldsymbol{\Lambda}\mathbf{V})^T + \boldsymbol{\Psi} = (\boldsymbol{\Lambda}\mathbf{V} + \mathbf{M}\mathbf{Q}^T)(\boldsymbol{\Lambda}\mathbf{V} + \mathbf{M}\mathbf{Q}^T)^T + (\boldsymbol{\Psi} - \mathbf{M}\mathbf{M}^T).$$

Equation (6) implies that no consistent estimator of  $\boldsymbol{\Psi}$  exists if  $\boldsymbol{\Lambda}$  fails to be of full column rank. This may induce corresponding multimodalities in the densities of  $\boldsymbol{\Psi}$  and  $\boldsymbol{\Lambda}$  [74], and is related to the choice of factor dimensionality and the possibility of retaining too many factors.

The FA model also copes with an inherent indeterminacy of the parameters, being: Rotational indeterminacy of the factor solution. Assume that  $\mathbf{R} \in \mathbb{R}^{m \times m}$  is an arbitrary nonsingular matrix. Returning to the implied covariance structure of the observed data, we then have

$$(7) \quad \boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi} = (\boldsymbol{\Lambda}\mathbf{R})[\mathbf{R}^{-1}\boldsymbol{\Phi}(\mathbf{R}^T)^{-1}](\boldsymbol{\Lambda}\mathbf{R})^T + \boldsymbol{\Psi},$$

implying that there is an infinite number of alternative matrices  $\boldsymbol{\Lambda}^\dagger = \boldsymbol{\Lambda}\mathbf{R}$  and  $\boldsymbol{\Phi}^\dagger = \mathbf{R}^{-1}\boldsymbol{\Phi}(\mathbf{R}^T)^{-1}$  that generate the same covariance structure  $\boldsymbol{\Sigma}$ . The operation  $\boldsymbol{\Lambda} \mapsto \boldsymbol{\Lambda}\mathbf{R}$  is termed ‘rotation’. Thus, in any solution,  $\boldsymbol{\Lambda}$  can be made to satisfy  $m^2$

additional conditions, which is naturally equivalent to the number of independent elements of  $\mathbf{R}$ .

From the above it is clear that any method of estimation requires at a minimum  $m^2$  restrictions on  $\mathbf{\Lambda}$  and  $\mathbf{\Phi}$ . The EFA tradition usually achieves this by requiring that  $\mathbf{\Phi} = \mathbf{I}_m$  and  $\mathbf{\Lambda}^T \mathbf{\Psi}^{-1} \mathbf{\Lambda}$  be diagonal accompanied by an order condition on the diagonal elements. These restrictions are arbitrary such that whence estimation is settled EFA traditionally endeavors on applying a rotation that satisfies certain criteria for interpretation purposes. Peeters [68] has given minimal conditions for the formulation of a rotationally unique confirmatory unrestricted factor model that cancels the need for post-hoc rotation.

#### APPENDIX: A PRIMER ON BAYESIAN STATISTICS

The Bayesian viewpoint is distinct from the classical approach to statistics. Let  $\Theta$  denote a model-specific collection of unknown parameters of continuous metric. The frequentist approach solely utilizes the likelihood of the observed data  $L(\Theta; \mathbf{X})$ , in that a retrospective evaluation is made of a certain statistic used to estimate  $\Theta$  over all possible  $\mathbf{X}$  values conditional on the true unknown  $\Theta$  which is deemed fixed. The Bayesian approach views  $\Theta$  as random. This allows for probability statements about  $\pi(\Theta|\mathbf{X})$ , the distribution of model parameters conditioned on the observed data. To provide the mentioned conditional probabilities, a joint probability function for  $\Theta$  and  $\mathbf{X}$  must be provided for. To this purpose a prior distribution  $\pi(\Theta)$  must be specified, which reflects the formalized knowledge or uncertainty about the parameters before observation of the data. Using a basic property of conditional probability known as Bayes' rule [80, 81], one obtains the posterior distribution as:

$$(8) \quad \pi(\Theta|\mathbf{X}) = \frac{L(\Theta; \mathbf{X})\pi(\Theta)}{\int L(\Theta; \mathbf{X})\pi(\Theta) d\Theta}.$$

Expression (8) encapsulates the core machinery of Bayesian statistics, whose flexibility has proven to extend to complex problems (consult, for example [82, 83]). The denominator in (8) is called the prior predictive density or marginal likelihood and is key in Bayesian model selection.

Let us shortly review Bayesian model selection for latent variable models. Let  $\vartheta$  denote latent data. For the factor model described in Section 'The Factor Analytic Model',  $\Theta = \{\mu, \mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{\Phi}\}$  and  $\vartheta = \Xi$ . Now, let  $g(\vartheta|\Theta)$  denote the density of latent data  $\vartheta$  given  $\Theta$  and assume that the complete data likelihood consists of  $L(\Theta, \vartheta; \mathbf{X})g(\vartheta|\Theta)$ . Suppose also that the prior  $\pi(\Theta)$  is available for the unknown model parameters  $\Theta$ . The marginal likelihood is then expressed as:

$$(9) \quad m(\mathbf{X}) = \int L(\Theta, \vartheta; \mathbf{X})\pi(\Theta)g(\vartheta|\Theta) d(\Theta, \vartheta).$$

The marginal likelihood expresses the likelihood of the data conditional on the model entertained. This quantity is of import in the construction of the Bayes factor, the main Bayesian model selection criterion. Suppose that  $S$  competing models  $M_s$  are under consideration, for  $s = 1, \dots, S$ . The Bayes factor of  $M_s$  to

$M_{s'}$  is then expressed as [84, 85]:

$$(10) \quad B_{ss'} = \frac{m_s(\mathbf{X})}{m_{s'}(\mathbf{X})} = \frac{\int L_s(\boldsymbol{\Theta}_s, \boldsymbol{\vartheta}_s; \mathbf{X}) \pi_s(\boldsymbol{\Theta}_s) g_s(\boldsymbol{\vartheta}_s | \boldsymbol{\Theta}_s) \partial(\boldsymbol{\Theta}_s, \boldsymbol{\vartheta}_s)}{\int L_{s'}(\boldsymbol{\Theta}_{s'}, \boldsymbol{\vartheta}_{s'}; \mathbf{X}) \pi_{s'}(\boldsymbol{\Theta}_{s'}) g_{s'}(\boldsymbol{\vartheta}_{s'} | \boldsymbol{\Theta}_{s'}) \partial(\boldsymbol{\Theta}_{s'}, \boldsymbol{\vartheta}_{s'})}.$$

The Bayes factor embodies the ratio of posterior odds to prior odds for the models under consideration. The expression in (10) resembles a likelihood ratio. But instead of evaluating the respective likelihoods at the maximum likelihood estimates, the parameters are integrated out with respect to the priors. The BF thus can be viewed as representing a ‘weighted’ likelihood ratio that provides a measure “of the evidence provided by the data in favor of one scientific theory, represented by a statistical model, as opposed to another” [66]. The Bayes factor behaves like a natural Occam’s razor, as model fit and complexity are accounted for in the marginal likelihood [86, 87]. For interpretation the quantity (10) can be referred to half-units on the  $\log_{10}$  scale (Appendix B of [85]) or one can consider  $2 \log_e B_{ss'}$  [66], which is on the same scale as likelihood ratio statistics. Another interpretational aid might be the posterior model probability, defined as [72]:

$$(11) \quad P(M_{s'} | \mathbf{X}) = \left( \sum_{s=1}^S \frac{p_s}{p_{s'}} \cdot B_{ss'} \right)^{-1}.$$

In (11)  $p_s$  denotes the prior probability one assigns to model  $M_s$  being best,  $\sum_{s=1}^S p_s = 1$ . The posterior model probability  $P(M_{s'} | \mathbf{X})$  gives the posterior probability, given the batch of models under consideration, that model  $M_{s'}$  is the correct model for the data at hand. A normalization of the Bayes factor ensues when letting  $p_s = S^{-1} \forall s$ .

The Bayes factor as a model selection criterion has several advantages (cf., [66, 73]): (i) It provides both a measure of evidence against a competing model and a measure of support for the alternative model; (ii) It will not by default favor the alternative model in (very) large samples; (iii) It allows one to take model uncertainty into account, thus providing a consistent quantity for the comparison of a multitude of competing models; (iv) It can handle the comparison of both nested and nonnested models. For an introduction to Bayesian statistics see [82, 83]. For Bayesian factor analysis see [73, 88]. For inequality-constrained-model selection for Bayesian FA as well as Bayesian treatments of both EFA and CFA, see [67].

#### APPENDIX: RESULTS FACTOR ANALYTIC APPROACH WEISS ET AL. [69]

Table 2 contains the results of the little jiffy approach to the data as given in the original study (Table 3 in [69]). A three-component solution was obtained in which the first component was interpreted as ‘obesity and glucose metabolism’, the second as ‘dyslipidemia’, and the third as ‘blood pressure’.

#### APPENDIX: REPRODUCED CORRELATION STRUCTURE

Table 3 contains the sample correlations, the correlations reproduced by the factor analysis, and the residual correlations (sample correlation minus reproduced correlation).

TABLE 2. Results little jiffy analysis on the data [69]

Components			Item
1	2	3	
.44	.13	.06	BMI
.09	.83	.04	$\log_e\{\text{trig.}\}$
-.13	-.82	.06	HDL chol.
.76	.27	.15	$\log_e\{\text{IR}\}$
.72	-.14	.07	$\log_e\{\text{GB}\}$
.67	.10	-.06	$\log_e\{\text{G2}\}$
.15	.09	.79	SBP
.01	.09	.83	DBP

## REFERENCES

- [1] Marañon G. Über Hypertonie und Zuckerkrankheit. Zentralblatt für Innere Medizin. 1922;43:169–176.
- [2] Kylin E. Studien über das Hypertonie-Hyperglykämie-Hyperurikämie Syndrom. Zentralblatt für Innere Medizin. 1923;44:105–112.
- [3] Reaven GM. Banting Lecture 1988. Role of insulin resistance in human disease. Diabetes. 1988;37:1595–1607.
- [4] Reaven GM. Role of insulin resistance in human disease (syndrome X): An expanded definition. Annual Review of Medicine. 1993;44:121–131.
- [5] Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. The Lancet. 2005;365:1415–1428.
- [6] Unwin N. The metabolic syndrome. Journal of the Royal Society of Medicine. 2006;99:457–462.
- [7] Wilkin TJ, Voss LD. Metabolic syndrome: Maladaptation to a modern world. Journal of the Royal Society of Medicine. 2004;97:511–520.
- [8] Miranda PJ, DeFronzo RA, Califf RM, Gruyton JR. Metabolic syndrome: Definition, pathophysiology, and mechanisms. American Heart Journal. 2005;149:33–45.
- [9] Grundy SM. Point: The metabolic syndrome still lives. Clinical Chemistry. 2005;51:1352–1354.
- [10] Kahn R, Buse J, Ferrannini E, Stern M. The metabolic syndrome: Time for a critical appraisal. Joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. Diabetes Care. 2005;28:2289–2304.
- [11] Reaven GM. Counterpoint: Just being alive is not good enough. Clinical Chemistry. 2005;51:1354–1357.
- [12] Reaven GM. The metabolic syndrome: Requiescat in Pace. Clinical Chemistry. 2005;51:931–938.
- [13] Blaha M, Elasy TA. Clinical use of the metabolic syndrome: Why the confusion? Clinical Diabetes. 2006;24:125–131.
- [14] Laakso M, Kovanen PT. Metabolic syndrome: To be or not to be? Annals of Medicine. 2006;38:32–33.
- [15] Venkat Narayan KM. The metabolic syndrome: Some second thoughts? Clinical Diabetes. 2006;24:38–39.
- [16] Alberti KGM, Zimmet PZ. Should we dump the metabolic syndrome?: No. British Medical Journal. 2008;336:641.
- [17] Gale EAM. Should we dump the metabolic syndrome?: Yes. British Medical Journal. 2008;336:640.
- [18] Tenenbaum A, Fisman E. “The metabolic syndrome... is dead”: These reports are an exaggeration. Cardiovascular Diabetology. 2011;10:11.

TABLE 3. Matrix containing observed (Pearson), reproduced, and residual correlations

	1	2	3	4	5	6	7	8
1 BMI								
observed	1.000							
reproduced	1.002							
residual	-.002							
2 $\log_e\{\text{trig.}\}$								
observed	.041	1.000						
reproduced	.101	.984						
residual	-.060	.016						
3 HDL chol.								
observed	-.143	-.423	1.000					
reproduced	-.111	-.459	.998					
residual	-.032	.036	.002					
4 $\log_e\{\text{IR}\}$								
observed	.314	.259	-.254	1.000				
reproduced	.291	.250	-.276	1.006				
residual	.023	.009	.022	-.006				
5 $\log_e\{\text{GB}\}$								
observed	.071	.051	-.066	.387	1.000			
reproduced	.161	.082	-.092	.384	1.004			
residual	-.090	-.031	.026	.003	-.004			
6 $\log_e\{\text{G2}\}$								
observed	.121	.197	-.105	.336	.233	1.000		
reproduced	.141	.143	-.157	.338	.183	.999		
residual	-.020	.054	.052	-.002	.050	.001		
7 SBP								
observed	.132	.065	-.028	.190	.104	.084	1.000	
reproduced	.089	.029	-.033	.212	.125	.100	.999	
residual	.043	.036	.005	-.022	-.021	-.016	.001	
8 DBP								
observed	-.013	-.016	.084	.093	.095	-.010	.332	1.000
reproduced	.047	-.056	.059	.111	.077	.048	.047	.999
residual	-.060	.004	.025	-.018	.018	-.058	.275	.001

- [19] World Health Organization. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Report of a WHO Consultation. Part 1: Diagnosis and Classification of Diabetes Mellitus. Geneva: WHO; 1999. WHO/NCD/NCS/99.2.
- [20] Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Journal of the American Medical Association. 2001;285:2486–2497.



- [21] Scuteri A, Najjar SS, Morrell CH, Lakatta EG. The metabolic syndrome in older individuals: Prevalence and prediction of cardiovascular events. The cardiovascular health study. *Diabetes Care*. 2005;28:882–887.
- [22] Liese AD, Mayer-Davis EJ, Haffner SM. Development of the multiple metabolic syndrome: An epidemiologic perspective. *Epidemiologic Reviews*. 1998;20:157–172.
- [23] Meigs JB. Invited Commentary: Insulin resistance syndrome? Syndrome X? Multiple metabolic syndrome? A syndrome at all? Factor analysis reveals patterns in the fabric of correlated metabolic risk factors. *American Journal of Epidemiology*. 2000;152:908–911.
- [24] Mulaik SA. *Foundations of Factor Analysis*. 2nd ed. Boca Raton [etc.]: Chapman & Hall/CRC; 2010.
- [25] Lawlor DA, Ebrahim S, May M, Smith GD. (Mis)use of factor analysis in the study of insulin resistance syndrome. *American Journal of Epidemiology*. 2004;159:1013–1018.
- [26] Edwards KL, Austin MA, Newman B, Mayer E, Krauss RM, Selby JV. Multivariate analysis of the insulin resistance syndrome in women. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 1994;14:1940–1945.
- [27] Pison G, Rousseeuw PJ, Croux C. Robust factor analysis. *Journal of Multivariate Analysis*. 2003;84:145–172.
- [28] Lee SE, Press SJ. Robustness of Bayesian factor analysis estimates. *Communications in Statistics – Theory and Methods*. 1998;27:1871–1893.
- [29] Kaiser HF. A second generation little jiffy. *Psychometrika*. 1970;35:401–415.
- [30] Jöreskog KG. Factor analysis and its extensions. In: Cudeck R, MacCallum RC, editors. *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007. p. 47–77.
- [31] Moustaki I. Factor analysis and latent structure of categorical and metric data. In: Cudeck R, MacCallum RC, editors. *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007. p. 293–313.
- [32] Borsboom D, Mellenbergh GJ, van Heerden J. The theoretical status of latent variables. *Psychological Review*. 2003;110:203–219.
- [33] Hanley AJG, Karter AJ, Festa A, D’Agostino R, Wagenknecht LE, Savage P, et al. Factor analysis of metabolic syndrome using directly measured insulin sensitivity: The Insulin Resistance Atherosclerosis Study. *Diabetes*. 2002;51:2642–2647.
- [34] Hanson RL, Imperatore G, Bennett PH, Knowler WC. Components of the “metabolic syndrome” and incidence of type 2 diabetes. *Diabetes*. 2002;51:3120–3127.
- [35] Ford ES. Factor analysis and defining the metabolic syndrome. *Ethnicity & Disease*. 2003;13:429–437.
- [36] Lambert M, Paradis G, O’Loughlin J, Delvin EE, Hanley JA, Levy E. Insulin resistance syndrome in a representative sample of children and adolescents from Quebec, Canada. *International Journal of Obesity*. 2004;28:833–841.
- [37] Oh JY, Hong YS, Sung YA, Barrett-Connor E. Prevalence and factor analysis of metabolic syndrome in an urban Korean population. *Diabetes Care*. 2004;27:2027–2032.
- [38] Ang LW, Ma S, Cutter J, Chew SK, Tan CE, Tai ES. The metabolic syndrome in Chinese, Malays, and Asian Indians: Factor analysis of data from the 1998 Singapore National Health Survey. *Diabetes Research and Clinical Practice*. 2005;67:53–62.
- [39] Ghosh A. Factor analysis of metabolic syndrome among the middle-aged Bengalee Hindu men of Calcutta, India. *Diabetes/Metabolism Research and Reviews*. 2005;21:58–64.
- [40] Corsetti JP, Rainwater DL, Moss AJ, Zareba W, Sparks CE. High lipoprotein-associated phospholipase A<sub>2</sub> is a risk factor for recurrent coronary events in postinfarction patients. *Clinical Chemistry*. 2006;52:1331–1338.
- [41] Mannucci E, Monami M, Rotella CM. How many components for the metabolic syndrome? Results of exploratory factor analysis in the FIBAR study. *Nutrition, Metabolism & Cardiovascular Diseases*. 2007;17:719–726.
- [42] Huang P, Kraja AT, Tang W, Hunt SC, North KE, Lewis CE, et al. Factor relationships of metabolic syndrome and echocardiographic phenotypes in the HyperGEN study. *Journal of Hypertension*. 2008;26:1360–1366.
- [43] Khader YS, Batieha A, Jaddou H, Batieha Z, El-Khateeb M, Ajlouni K. Factor analysis of cardiometabolic risk factors clustering in children and adolescents. *Metabolic Syndrome and Related Disorders*. 2011;9:151–156.

- [44] Widaman KF. Common factors versus components: Principals and principles, errors and misconceptions. In: Cudeck R, MacCallum RC, editors. *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007. p. 177–203.
- [45] Jolliffe IT. *Principal Component Analysis*. 2nd ed. New York [etc.]: Springer; 2002.
- [46] Guttman L. Some necessary conditions for common factor analysis. *Psychometrika*. 1954;19:149–162.
- [47] Kaiser HF. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*. 1960;20:141–151.
- [48] Yeomans KA, Golder PA. The Guttman-Kaiser criterion as a predictor of the number of common factors. *The Statistician*. 1982;31:221–229.
- [49] Cliff N. The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin*. 1988;103:276–279.
- [50] Kaiser HF. The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psychometrika*. 1958;23:187–200.
- [51] Preacher KJ, MacCallum RC. Repairing Tom Swift’s Electric Factor Analysis Machine. *Understanding Statistics*. 2003;2:13–43.
- [52] Shen BJ, Todaro JF, Niaura R, McCaffery JM, Zhang J, Spiro A, et al. Are metabolic risk factors one unified syndrome? Modeling the structure of the metabolic syndrome X. *American Journal of Epidemiology*. 2003;157:701–711.
- [53] Pladevall M, Singal B, Williams LK, Brotons C, Guyer H, Sadurni J, et al. A single factor underlies the metabolic syndrome: A confirmatory factor analysis. *Diabetes Care*. 2006;29:113–122.
- [54] Shah S, Novak S, Stapleton LM. Evaluation and comparison of models of metabolic syndrome using confirmatory factor analysis. *European Journal of Epidemiology*. 2006;21:343–349.
- [55] Chaoyang L, Ford ES. Is there a single underlying factor for the metabolic syndrome in adolescents? *Diabetes Care*. 2007;30:1556–1561.
- [56] Goodman E, Li C, Tu YK, Ford ES, Sun SS, Huang TTK. Stability of the factor structure of the metabolic syndrome across pubertal development: Confirmatory factor analyses of three alternative models. *The Journal of Pediatrics*. 2009;155:S5.e1–S5.e8.
- [57] Boronat M, Saavedra P, Varillas VF, Nóvoa FJ. Use of confirmatory factor analysis for the identification of new components of the metabolic syndrome: The role of plasminogen activator inhibitor-1 and Haemoglobin A1c. *Nutrition, Metabolism & Cardiovascular Diseases*. 2009;19:271–276.
- [58] Solera-Martínez M, López-Martínez S, Sánchez-López M, Moya-Martínez P, Notario-Pacheco B, Arias-Palencia N, et al. Validity of a single-factor model underlying the metabolic syndrome in young adults: Confirmatory factor analysis. *Revista Española de Cardología*. 2011;64:379–384.
- [59] Hoyle RH, Duvall JL. Determining the Number of Factors in Exploratory and Confirmatory Factor Analysis. In: Kaplan D, editor. *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks, CA: SAGE; 2004. p. 301–315.
- [60] Ferrando PJ, Lorenzo-Seva U. Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica*. 2000;21:301–323.
- [61] van Prooijen JW, van der Kloot WA. Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*. 2001;61:777–792.
- [62] Marsh HW, Lüdtke O, Nagengast B, Morin AJS, Von Davier M. Why item parcels are (almost) never appropriate: Two wrongs do not make a right – camouflaging misspecification with item parcels in CFA models. *Psychological Methods*. 2013;18:257–284.
- [63] Little TD, Rhemtulla M, Gibson K, Schoemann AM. Why the items versus parcels controversy needn’t be one. *Psychological Methods*. 2013;18:285–300.
- [64] Platt JR. Strong inference. *Science*. 1964;146:347–353.
- [65] Dunson DB. Commentary: Practical advantages of Bayesian analysis of epidemiologic data. *American Journal of Epidemiology*. 2001;153:1222–1226.
- [66] Kass RE, Raftery AE. Bayes factors. *Journal of the American Statistical Association*. 1995;90:773–795.

- [67] Peeters CFW. Bayesian Exploratory and Confirmatory Factor Analysis: Perspectives on Constrained-Model Selection. Unpublished PhD thesis, Dept. of Methodology & Statistics, Utrecht University; 2012.
- [68] Peeters CFW. Rotational uniqueness conditions under oblique factor correlation metric. *Psychometrika*. 2012;77:288–292.
- [69] Weiss R, Dziura J, Burgert TS, Tamborlane WV, Taksali SE, Yeckel CW, et al. Obesity and the metabolic syndrome in children and adolescents. *The New England Journal of Medicine*. 2004;350:2362–2374.
- [70] Einhorn D, Reaven GM, Cobin RH, Ford ES, Ganda OP, Handelsman Y, et al. American College of Endocrinology position statement on the insulin resistance syndrome. *Endocrine Practice*. 2003;9:237–252.
- [71] Chib S. Marginal Likelihood from the Gibbs output. *Journal of the American Statistical Association*. 1995;90:1313–1321.
- [72] Berger JO, Pericchi LR. The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*. 1996;91:109–122.
- [73] Lee SY. Structural Equation Modeling: A Bayesian Approach. West Sussex: John Wiley and Sons Ltd.; 2007.
- [74] Lopes HF, West M. Bayesian model assessment in factor analysis. *Statistica Sinica*. 2004;14:41–67.
- [75] Hayashi K, Bentler PM, Yuan KH. On the likelihood ratio test for the number of factors in EFA. *Structural Equation Modeling*. 2007;14:505–526.
- [76] Klugkist I, Hoijsink H. The Bayes factor for inequality and about equality constrained models. *Computational Statistics & Data Analysis*. 2007;51:6367–6379.
- [77] Bekker PA, Merckens A, Wansbeek TJ. Identification, Equivalent Models, and Computer Algebra. Boston [etc.]: Academic Press; 1994.
- [78] Anderson TW, Rubin H. Statistical inference in factor analysis. In: Neyman J, editor. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. vol. V. Berkeley: University of California Press; 1956. p. 111–150.
- [79] Geweke JF, Singleton KJ. Interpreting the likelihood ratio statistic in factor models when sample size is small. *Journal of the American Statistical Association*. 1980;75:133–137.
- [80] Bayes T. An essay towards solving a problem in the doctrine of chances. *Biometrika*. 1958;45:293–315. Reproduced from: *The Philosophical Transactions (1763)*, **53**, 370–418; with a foreword by G.A. Barnard.
- [81] Laplace PS. Memoir on the probability of the causes of events. *Statistical Science*. 1986;1:364–378. Translated by S.M. Stigler, and reproduced from: *Mémoires de Mathématique et de Physique, Présentés à l'Académie Royale des Sciences, par divers Savans & lus dans ses Assemblées, Tome Sixieme (1774)*, 621–656.
- [82] Press SJ. Subjective and Objective Bayesian Statistics: Principles, Models, and Applications. 2nd ed. New York: Wiley; 2003.
- [83] Gelman A, Carlin BP, Stern HS, Rubin DB. *Bayesian Data Analysis*. Boca Raton: Chapman & Hall/CRC; 2004.
- [84] Jeffreys H. Some tests of significance, treated by the theory of probability. *Proceedings of the Cambridge Philosophy Society*. 1935;31:203–222.
- [85] Jeffreys H. *Theory of Probability*. 3rd ed. Oxford: Oxford University Press; 1961.
- [86] Spiegelhalter DJ, Smith AFM. Bayes factor for linear and log-Linear models with vague prior information. *Journal of the Royal Statistical Society, Series B*. 1982;44:377–387.
- [87] Jefferys W, Berger JO. Ockham's razor and Bayesian analysis. *American Scientist*. 1992;80:64–72.
- [88] Muthén B, Asparouhov T. Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*. 2012;17:313–335.
- [89] PubMed [Bethesda (MD): National Library of Medicine (US)] Available from: <http://www.ncbi.nlm.nih.gov/pubmed/>;
- [90] Penno G, Miccoli R, Pucci L, Del Prato S. The metabolic syndrome: Beyond the insulin resistance syndrome. *Pharmacological Research*. 2006;53:457–468.
- [91] Ford ES, Li C. Defining the metabolic syndrome in children and adolescents: Will the real definition please stand up? *The Journal of Pediatrics*. 2008;152:160–164.
- [92] Bollen KA. *Structural Equations with Latent Variables*. New York: Wiley; 1989.

- [93] Chen W, Srinivasan SR, Elkasabany A, Berenson GS. Cardiovascular risk factors clustering features of insulin resistance syndrome (Syndrome X) in a biracial (Black-White) population of children, adolescents, and young adults: the Bogalusa Heart Study. *American Journal of Epidemiology*. 1999;150:667–674.
- [94] Chen W, Srinivasan SR, Berenson GS. Plasma renin activity and insulin resistance in African American and white children: the Bogalusa Heart Study. *American Journal of Hypertension*. 2001;14:212–217.
- [95] Dwyer T, Blizzard L, Venn A, Stankovich JM, Ponsonby AL, Morley R. Syndrome X in 8-y-old Australian children: Stronger associations with current body fatness than with infant size or growth. *International Journal of Obesity*. 2002;26:1301–1309.
- [96] Moreno LA, Pineda I, Rodríguez G, Fleta J, Giner A, Juste MG, et al. Leptin and metabolic syndrome in obese and non-obese children. *Hormone and Metabolic Research*. 2002;34:394–399.
- [97] Park HS, Lee MS, Park JY. Leptin and the metabolic syndrome in Korean adolescents: Factor analysis. *Pediatrics International*. 2004;46:697–703.
- [98] Goodman E, Dolan LM, Morrison JA, Daniels SR. Factor analysis of clustered cardiovascular risks in adolescence: obesity is the predominant correlate of risk among youth. *Circulation*. 2005;111:1970–1977.
- [99] Retnakaran R, Zinman B, Connelly PW, Harris SB, J HA. Nontraditional cardiovascular risk factors in pediatric metabolic syndrome. *The Journal of Pediatrics*. 2006;148:176–182.
- [100] Ghosh A. Factor analysis of risk variables associated with metabolic syndrome in Asian Indian adolescents. *American Journal of Human Biology*. 2007;19:34–40.
- [101] Goodman E, Daniels SR, Meigs JB, Dolan LM. Instability in the diagnosis of metabolic syndrome in adolescents. *Circulation*. 2007;115:2316–2322.
- [102] Kelishadi R, Ardalan G, Adeli K, Motaghian M, Majdzadeh R, Mahmood-Arabi MS, et al. Factor analysis of cardiovascular risk clustering in pediatric metabolic syndrome: CASPIAN study. *Annals of Nutrition & Metabolism*. 2007;51:208–215.
- [103] Li C, Ford ES. Is there a single underlying factor for the metabolic syndrome in adolescents? A confirmatory factor analysis. *Diabetes Care*. 2007;30:1556–1561.
- [104] Ng VW, Kong AP, Choi KC, Ozaki R, Wong GW, So WY, et al. BMI and waist circumference in predicting cardiovascular risk factor clustering in Chinese adolescents. *Obesity*. 2007;15:494–503.
- [105] Ramachandran A, Snehalatha C, Yamuna A, Murugesan N, Narayan KM. Insulin resistance and clustering of cardiometabolic risk factors in urban teenagers in southern India. *Diabetes Care*. 2007;30:1828–1833.
- [106] Mayer-Davis EJ, Ma B, Lawson A, D’Agostino RB, Liese AD, Bell RA, et al. Cardiovascular disease risk factors in youth with type 1 and type 2 diabetes: Implications of a factor analysis of clustering. *Metabolic Syndrome and Related Disorders*. 2009;7:89–95.
- [107] Noto D, Niglio T, Cefalù AB, Martino E, Fayer F, Mina M, et al. Obesity and the metabolic syndrome in a student cohort from Southern Italy. *Nutrition, Metabolism & Cardiovascular Diseases*. 2009;19:620–625.
- [108] Ghasemi A, Zahediasl S, Azizi F. Nitric oxide and clustering of metabolic syndrome components in pediatrics. *European Journal of Epidemiology*. 2010;25:45–53.
- [109] Martínez-Vizcaíno V, Martínez MS, Aguilar FS, Martínez SS, Gutiérrez RF, López MS, et al. Validity of a single-factor model underlying the metabolic syndrome in children: A confirmatory factor analysis. *Diabetes Care*. 2010;33:1370–1372.
- [110] Martínez-Vizcaíno V, B OF, Solera-Martínez M, Ruiz JR, Labayen I, Eensoo D, et al. Stability of the factorial structure of metabolic syndrome from childhood to adolescence: A 6-year follow-up study. *Cardiovascular Diabetology*. 2011;10:81.
- [111] Olza J, Gil-Campos M, Leis R, Bueno G, Aguilera CM, Valle M, et al. Presence of the metabolic syndrome in obese children at prepubertal age. *Annals of Nutrition & Metabolism*. 2011;58:343–350.
- [112] Gurka MJ, Ice CL, Sun SS, Deboer MD. A confirmatory factor analysis of the metabolic syndrome in adolescents: An examination of sex and racial/ethnic differences. *Cardiovascular Diabetology*. 2012;11:128.
- [113] Hong TK, Trang NH, Dibley MJ. Prevalence of metabolic syndrome and factor analysis of cardiovascular risk clustering among adolescents in Ho Chi Minh City, Vietnam. *Preventive Medicine*. 2012;55:409–411.

- [114] Suchday S, Bellehsen M, Friedberg JP, Almeida M, Kaplan E. Clustering of cardiac risk factors associated with the metabolic syndrome and associations with psychosocial distress in a young Asian Indian population. *Journal of Behavioral Medicine*. 2013;doi: 10.1007/s10865-013-9521-x:[Epub ahead of print].
- [115] Wang Q, Yin J, Xu L, Cheng H, Zhao X, Xiang H, et al. Prevalence of metabolic syndrome in a cohort of Chinese schoolchildren: Comparison of two definitions and assessment of adipokines as components by factor analysis. *BMC Public Health*. 2013;13:249.
- [116] Edwards KL, Newman B, Mayer E, Selby JV, Krauss RM, Austin MA. Heritability of factor of the insulin resistance syndrome in women twins. *Genetic Epidemiology*. 1997;14:241–253.
- [117] Donahue RP, Bean JA, Donahue RD, Goldberg RB, Prineas RJ. Does insulin resistance unite the separate components of the insulin resistance syndrome? Evidence from the Miami Community Health Study. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 1997;17:2413–2417.
- [118] Meigs JB, D’Agostino RB, Wilson PW, Cupples LA, Nathan DM, Singer DE. Risk variable clustering in the insulin resistance syndrome: The Framingham Offspring Study. *Diabetes*. 1997;46:1594–1600.
- [119] Edwards KL, Burchfiel CM, Sharp DS, Curb JD, Rodriguez BL, Fujimoto WY, et al. Factors of the insulin resistance syndrome in nondiabetic and diabetic elderly Japanese-American men. *American Journal of Epidemiology*. 1998;147:441–447.
- [120] Gray RS, Fabsitz RR, Cowan LD, Lee ET, Howard BV, Savage PJ. Risk factor clustering in the insulin resistance syndrome: The Strong Heart Study. *American Journal of Epidemiology*. 1998;148:869–878.
- [121] Leyva F, Godsland IF, Ghatei M, Proudler AJ, Aldis S, Walton C, et al. Hyperleptinemia as a component of a metabolic syndrome of cardiovascular risk. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 1998;18:928–933.
- [122] Chen CH, Lin KC, Tsai ST, Chou P. Different association of hypertension and insulin-related metabolic syndrome between men and women in 8437 nondiabetic Chinese. *American Journal of Hypertension*. 2000;13:846–853.
- [123] Sakkinen PA, Wahl P, Cushman M, Lewis MR, Tracy RP. Clustering of procoagulation, inflammation, and fibrinolysis variables with metabolic factors in insulin resistance syndrome. *American Journal of Epidemiology*. 2000;152:897–907.
- [124] Snehalatha C, Sivasankari S, Satyavani K, Vijay V, Ramachandran A. Insulin resistance alone does not explain the clustering of cardiovascular risk factors in southern India. *Diabetic Medicine*. 2000;17:152–157.
- [125] Anderson PJ, Critchley JAJH, Chan JCN, Cockram CS, Lee ZSK, Thomas GN, et al. Factor analysis of the metabolic syndrome: Obesity vs insulin resistance as the central abnormality. *International Journal of Obesity*. 2001;25:1782–1788.
- [126] Hodge AM, Boyko EJ, de Courten M, Zimmet PZ, Chitson P, Tuomilehto J, et al. Leptin and other components of the metabolic syndrome in Mauritius – a factor analysis. *International Journal of Obesity*. 2001;25:126–131.
- [127] Shmulewitz D, Auerbach SB, Lehner T, Blundell ML, Winick JD, Youngman LD, et al. Epidemiology and factor analysis of obesity, type II diabetes, hypertension, and dyslipidemia (Syndrome X) on the island of Kosrae, Federated States of Micronesia. *Human Heredity*. 2001;51:8–19.
- [128] Adami GF, Civalleri D, Cella F, Marinari G, Camerini G, Papadia F, et al. Relationships of serum leptin to clinical and anthropometric findings in obese patients. *Obesity Surgery*. 2002;12:623–627.
- [129] Arya R, Blangero J, Williams K, Almasy L, Dyer TD, Leach RJ, et al. Factors of insulin resistance syndrome – Related phenotypes are linked to genetic locations on chromosomes 6 and 7 in nondiabetic Mexican-Americans. *Diabetes*. 2002;51:841–847.
- [130] Kue Young T, Chateau D, Zhang M. Factor analysis of ethnic variation in the multiple metabolic (insulin resistance) syndrome in three Canadian populations. *American Journal of Human Biology*. 2002;14:649–658.
- [131] O’Dea K, Rowley KG. Macrovascular disease risk factors and insulin resistance in Aboriginal and Torres Strait Islander people. *Journal of Diabetes and Its Complications*. 2002;16:9–16.
- [132] Choi KM, Lee J, Kim KB, Kim DR, Kim SK, Shin DH, et al. Factor analysis of the metabolic syndrome among elderly Koreans – the South-West Seoul Study. *Diabetic Medicine*. 2003;20:99–104.

- [133] Hanley AJ, Connelly PW, Harris SB, Zinman B. Adiponectin in a native Canadian population experiencing rapid epidemiological transition. *Diabetes Care*. 2003;26:3219–3225.
- [134] Howard BV, Criqui MH, Curb DJ, Rodabough R, Safford MM, Santoro N, et al. Risk factor clustering in the insulin resistance syndrome and its relationship to cardiovascular disease in postmenopausal White, Black, Hispanic, and Asian/Pacific Islander women. *Metabolism*. 2003;52:362–371.
- [135] Novak S, Stapleton LM, Litaker JR, Lawson KA. A confirmatory factor analysis evaluation of the coronary heart disease risk factors of metabolic syndrome with emphasis on the insulin resistance factor. *Diabetes, Obesity and Metabolism*. 2003;5:388–396.
- [136] Tang W, Miller MB, Rich SS, North KE, Pankow JS, Borecki IB, et al. Linkage analysis of a composite factor for the multiple metabolic syndrome: The National Heart, Lung, and Blood Institute Family Heart Study. *Diabetes*. 2003;52:2840–2847.
- [137] Cai G, Cole SA, Freeland-Graves JH, MacCluer JW, Blangero J, Comuzzie AG. Principal component for metabolic syndrome risk maps to chromosome 4p in Mexican Americans: The San Antonio Family Heart Study. *Human Biology*. 2004;76:651–665.
- [138] Jones LM, Legge M, Goulding A. Factor analysis of the metabolic syndrome in spinal cord-injured men. *Metabolism*. 2004;53:1372–1377.
- [139] Nelson TL, Palmer RF, Pedersen NL. The metabolic syndrome mediates the relationship between cynical hostility and cardiovascular disease. *Experimental Aging Research*. 2004;30:163–177.
- [140] Wang JJ, Qiao Q, Miettinen ME, Lappalainen J, Hu G, Tuomilehto J. The metabolic syndrome defined by factor analysis and incident type 2 diabetes in a Chinese population with high postprandial glucose. *Diabetes Care* 2004 Oct;27(10):2429–37. 2004;27:2429–2437.
- [141] Lin H, Boden-Albala B, Juo SH, Park N, Rundek T, Sacco RL. Heritabilities of the metabolic syndrome and its components in the Northern Manhattan Family Study. *Diabetologia*. 2005;48:2006–2012.
- [142] Mohan V, Deepa R, Pradeepa R, Vimalaswaran KS, Mohan A, Velmurugan K, et al. Association of low adiponectin levels with the metabolic syndrome – the Chennai Urban Rural Epidemiology Study (CURES-4). *Metabolism - Clinical and Experimental*. 2005;54:476–481.
- [143] Liou TL, Lin MW, Hsiao LC, Tsai TT, Chan WL, Ho LT, et al. Is hyperuricemia another facet of the metabolic syndrome? *Journal of the Chinese Medical Association*. 2006;69:104–109.
- [144] Shen BJ, Goldberg RB, Llabre MM, Schneiderman N. Is the factor structure of the metabolic syndrome comparable between men and women and across three ethnic groups: the Miami Community Health Study. *Annals of Epidemiology*. 2006;16:131–137.
- [145] Zanolini ME, Tosi F, Zoppini G, Castello R, Spiazzi G, Dorizzi R, et al. Clustering of cardiovascular risk factors associated with the insulin resistance syndrome: Assessment by principal component analysis in young hyperandrogenic women. *Diabetes Care*. 2006;29:372–378.
- [146] Chien KL, Hsu HC, Chen WJ, Chen MF, Su TC, Lee YT. Familial aggregation of metabolic syndrome among the Chinese: Report from the Chin-Shan Community Family Study. *Diabetes Research and Clinical Practice*. 2007;76:418–424.
- [147] Razak F, Anand SS, Shannon H, Vuksan V, Davis B, Jacobs R, et al. Defining obesity cut points in a multiethnic population. *Circulation*. 2007;115:2111–2118.
- [148] Ghosh A. Comparison of risk variables associated with the metabolic syndrome in pre- and postmenopausal Bengalee women. *Cardiovascular Journal of Africa*. 2008;19:183–187.
- [149] Lafortuna CL, Adorni F, Agosti F, Sartorio A. Factor analysis of metabolic syndrome components in obese women. *Nutrition, Metabolism & Cardiovascular Diseases*. 2008;18:233–241.
- [150] Reimann M, Schutte AE, Malan L, Huisman HW, Malan NT. Hyperuricaemia is an independent factor for the metabolic syndrome in a sub-Saharan African population: A factor analysis. *Atherosclerosis*. 2008;197:638–645.
- [151] Wu CZ, Lin JD, Li JC, Hsiao FC, Hsieh CH, Kuo SW, et al. Factor analysis of metabolic syndrome using direct measurement of insulin resistance in Chinese with different degrees of glucose tolerance. *Indian Journal of Medical Research*. 2008;127:336–343.
- [152] Leone N, Courbon D, Thomas F, Bean K, Jégo B, Leynaert B, et al. Lung function impairment and metabolic syndrome: The critical role of abdominal obesity. *American Journal of Respiratory and Critical Care Medicine*. 2009;179:509–516.
- [153] Öhrvik J, Hedberg P, Jonason T, Lönnberg I, Nilsson G. Factor analysis of the individual components of the metabolic syndrome among elderly identifies two factors with different

- survival patterns – A population-based study. *Metabolic Syndrome and Related Disorders*. 2009;7:171–177.
- [154] Chang JW, Ou HY, Chen HL, Guo HR, Liao PC, Lee CC. Interrelationship between exposure to PCDD/Fs and hypertension in metabolic syndrome in Taiwanese living near a highly contaminated area. *Chemosphere*. 2010;81:1027–1032.
  - [155] Esteghamati A, Zandieh A, Khalilzadeh O, Morteza A, Meysamie A, Nakhjavani M, et al. Clustering of leptin and physical activity with components of metabolic syndrome in Iranian population: an exploratory factor analysis. *Endocrine*. 2010;38:206–213.
  - [156] Marsland AL, McCaffery JM, Muldoon MF, Manuck SB. Systemic inflammation and the metabolic syndrome among middle-aged community volunteers. *Metabolism - Clinical and Experimental*. 2010;59:1801–1809.
  - [157] McCaffery JM, Marsland AL, Strohacker K, Muldoon MF, Manuck SB. Factor structure underlying components of allostatic load. *PLoS One*. 2012;7:e47246.
  - [158] Barbosa-Leiker C, Wright BR, Burns GL, Parks CD, Strand PS. Longitudinal measurement invariance of the metabolic syndrome: Is the assessment of the metabolic syndrome stable over time? *Annals of Epidemiology*. 2011;21:111–117.
  - [159] Meshkani R, Zargari M, Larijani B. The relationship between uric acid and metabolic syndrome in normal glucose tolerance and normal fasting glucose subjects. *Acta Diabetologica*. 2011;48:79–88.
  - [160] Azimi-Nezhad M, Herbeth B, Siest G, Dadé S, Ndiaye NC, Esmaily H, et al. High prevalence of metabolic syndrome in Iran in comparison with France: What are the components that explain this? *Metabolic Syndrome and Related Disorders*. 2012;10:181–188.
  - [161] Esteghamati A, Zandieh A, Esteghamati A, Sadaghiani MS, Zandieh B, Rezaeitabar E, et al. Apolipoproteins A-I and B as components of metabolic syndrome with respect to diabetes status: A factor analysis. *Metabolic Syndrome and Related Disorders*. 2012;10:280–285.
  - [162] Stevenson JE, Wright BR, Boydston AS. The metabolic syndrome and coronary artery disease: A structural equation modeling approach suggestive of a common underlying pathophysiology. *Metabolism - Clinical and Experimental*. 2012;61:1582–1588.
  - [163] Chirinos DA, Medina-Lezama J, Arguelles W, Goldberg R, Schneiderman N, Khan Z, et al. Metabolic syndrome as an underlying disease entity and its relationship to subclinical atherosclerosis in Andean Hispanics. *Metabolic Syndrome and Related Disorders*. 2013;doi: 10.1089/met.2013.0092:[Epub ahead of print].
  - [164] Dussault-Belanger F, Cohen AA, Hivert MF, Courteau J, Vanasse A. Validating metabolic syndrome through principal component analysis in a medically diverse, realistic cohort. *Metabolic Syndrome and Related Disorders*. 2013;11:21–28.
  - [165] Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika*. 1965;30:179–185.
  - [166] Gómez-Marcos MA, Patino-Alonso MC, Recio-Rodríguez JI, Antón-Alvarez J, Cabrejas-Sánchez A, Fernandez-Alonso C, et al. Confirmatory factor analysis to assess the measure of adiposity that best fits the diagnosis of metabolic syndrome and relationship to physical activity in adults. *European Journal of Nutrition*. 2013;52:1451–1459.
  - [167] Gurka MJ, Lilly CL, Oliver MN, Deboer MD. An examination of sex and racial/ethnic differences in the metabolic syndrome among adults: A confirmatory factor analysis and a resulting continuous severity score. *Metabolism - Clinical and Experimental*. 2013;doi: 10.1016/j.metabol.2013.10.006.:[Epub ahead of print].
  - [168] Huo D, Wang W, Li X, Gao Q, Wu L, Luo Y, et al. Evaluation of two single-factor models of metabolic syndrome: A confirmatory factor analysis for an adult population in Beijing. *Lipids in Health and Disease*. 2013;12:61.
  - [169] Nilsson G, Hedberg P, Öhrvik J. Inflammation and the metabolic syndrome: Clustering and impact on survival in a Swedish community-based cohort of 75 year olds. *Metabolic Syndrome and Related Disorders*. 2013;11:92–101.
  - [170] Sabanayagam C, Teo BW, Tai ES, Jafar TH, Wong TY. Ethnic variation in the impact of metabolic syndrome components and chronic kidney disease. *Maturitas*. 2013;74:369–374.
  - [171] Smits MM, Woudstra P, Utzschneider KM, Tong J, Gerchman F, Faulenbach M, et al. Adipocytokines as features of the metabolic syndrome determined using confirmatory factor analysis. *Annals of Epidemiology*. 2013;23:415–421.

(Carel F.W. Peeters) DEPT. OF EPIDEMIOLOGY & BIostatISTICS, VU UNIVERSITY MEDICAL CENTER AMSTERDAM, AMSTERDAM, THE NETHERLANDS

*E-mail address:* `cf.peeters@vumc.nl`

(James Dziura) YALE CENTER FOR ANALYTICAL SCIENCES, YALE SCHOOL OF PUBLIC HEALTH, NEW HAVEN, CN, USA

*E-mail address:* `james.dziura@yale.edu`

(Floryt van Wesel) DEPT. OF METHODOLOGY & STATISTICS, UTRECHT UNIVERSITY, UTRECHT, THE NETHERLANDS

*E-mail address:* `F.vanWesel@uu.nl`



## SUPPLEMENTARY MATERIAL

Tables S1 and S2 below list overviews of published studies on the MBS employing factor analytic techniques. Table S1 lists studies among the child and adolescent cohort. Table S2 lists studies among subjects other than children and adolescents. Studies were found using the PubMed search engine [89] by pairing the search term “factor analysis” with each in {“metabolic syndrome”, “insulin resistance syndrome”, and “syndrome X”} (the names that are regularly used to refer to the MBS). By usage of these terms English-language publications were sought with (online) publication dates from 1994 (year of publication of the seminal article by [26]) to January 2014. See [90, 91] and the online supplement to [53] for additional (complementary) overview tables. Please consider the remarks below for full understanding of Tables S1 and S2.

**Eligibility criteria.** Included studies focus their factor analytic efforts, as in the main text, on variables of a metabolic nature. Studies that focus mainly on hemostatic, inflammatory, lifestyle, or diet variables were excluded. Studies that provided no information on their factor analytic procedure (extraction, retention, and rotation) were also excluded. In addition, studies that considered binary or categorical variables in standard factor models were also not considered.

**Data extraction.** If a study that focuses on the mentioned variables for exclusion in addition ran a factor analytic procedure on the traditional metabolic variables, then only the results of the analysis on the metabolic variables are reported. When study and validation cohorts were available in a certain study, only the study cohort is reported when results comply. Also, when FA was performed on various (sub)sets of variables, we chose to report the FA on the combination of items most in concurrence with the data (re-)analyzed in this study.

When representing EFA efforts (including PCA) the description of the factors is based on the loading cut-offs chosen by the authors of the original studies. When representing CFA efforts in the tables below, we focus on the factor structure, not on other model details such as correlated error variables. When a higher-order factor model is utilized in a certain study only the first-order factors are described (the second-order factor is always termed ‘MBS’). It is however explicitly indicated when a higher-order model is fitted. In addition, we only report final models when describing CFA efforts. Details such as usage of modification indices can be found in the studies themselves.

The assessment of the stability of the factor analytic model is generally known as the issue of ‘measurement invariance’. This issue arises in longitudinal data (is the model stable over the measured time-points) or when the total sample is considered to consist of subgroups (is the model stable over all subgroups that make up the total sample). Invariance can be studied on the level of the model structure (number of factors and factor structure) or, given a certain model structure, on the level of the parameter values. A formal approach would be to perform multiple group analyses, possibly paired with a hierarchy of invariance tests (see e.g., [92]). This formal approach is strongly tied to CFA. An informal approach, usually performed in EFA-type analyses, would be to perform separate EFA’s (or PCA’s) on the data from the respective subgroups or time-points and to assess loosely if the model is

stable. In the tables below we use ‘assessment of measurement invariance’ to refer to the more formal approach while we use ‘subgroup analyzes’ to refer to the more informal approach. In the tables below it is explicitly indicated if, within a certain study, factor analytic results differ over the various subgroups or time-points. If the model structure is considered invariant (in the original study) this will be designated with ‘structure considered invariant’. If model structure and parameter values are considered invariant (in the original study) this will be designated with ‘model considered invariant’.

**Reading the tables.** In the tables below ‘approach’ refers to the factor analytic approach taken and ‘factors’ refers to either common factors (in the case of a true factor analytic approach) or principal components (in case of a PCA approach). The sample size given is the size of the sample included in the factor analytic efforts. Characteristics refer to the characteristics of the sample included in the FA.

Note that in the description of factors the ordering is of import in the PCA approach and in an EFA approach with certain rotation criteria. The factors then can be understood as being in descending order of (percent of) variance explained (in the observed variables). In a CFA approach, the stated ordering of factors is arbitrary. The numbering in this latter case is solely to convey the number of factors modelled.

The reported studies are ordered according to year of publication. While the metabolic variables may at first seem to differ widely between the included studies, one may note that in many cases surrogate measures are used. For example, often the apolipoproteins B and A-I are used as surrogates for the LDL and HDL cholesterol fractions, respectively. Note also that the measures for insulin resistance can vary over studies, e.g., fasting and postload insulin, HOMA-IR, or the intravenous glucose tolerance test are some of the measures that are oft-used. In addition, the tables use ‘/’ to designate a ratio, e.g., trig./HDL chol. indicates the triglycerides over HDL cholesterol ratio.

The tables utilize, next to the abbreviations used in the main text, the following additional abbreviations:

: %BF = percent body fat; 17HP = 17-hydroxyprogesterone; Adi = adiponectin; Apo A-I = apolipoprotein A-I; Apo B = apolipoprotein B; AST = abdominal skinfold thickness; BW = body weight; CASPIAN = Childhood and Adolescence Surveillance and Prevention of Adult Non-Communicable Disease; FFFA = fasting free fatty acids; FG = fasting glucose; FI = fasting insulin; fib. = fibrinogen; FT = free testosterone; HbA1c = glycated hemoglobin; HC = hip circumference; HDL<sub>2</sub> = high density lipoprotein 2 cholesterol; HDL tot. = total high density lipoprotein cholesterol; IAF = (CT-measured) intra-abdominal fat area; IAI = Istituto Auxologico Italiano; IMGD = insulin-mediated glucose disposal; IS = insulin sensitivity; IVGTT = intravenous glucose tolerance test; LDL chol. = high density lipoprotein cholesterol; LDL-PPD = low-density lipoprotein peak particle diameter; LR = likelihood ratio; MAP = mean arterial pressure; ML = maximum likelihood; NFG = non fasting glucose; NHLBI = National Heart, Lung, and Blood Institute; NO<sub>x</sub> = nitric oxide metabolites; PAI-1 = plasminogen activator inhibitor-1; PAL = physical activity level

(minutes/week); PCDD/Fs = polychlorinated dibenzo-p-dioxin, dibenzofurans persistent organic pollutants; PFFA = postload free fatty acids; PG = postload glucose; PI = postload insulin; PIn = ponderal index; RA = renin activity; SSK = subscapular skinfold; SSPG = steady state plasma glucose; S:T = subscapular to triceps; SuRFNCD = Surveys of Risk Factors of Non-Communicable Diseases; TC = total cholesterol; TER = trunk extremity ratio; TFM = trunk fat mass; tria. = triacylglycerols; TSI = Torres Strait Islander; TTTS = trunk-to-total skinfolds; UA = uric acid; U:C = urinary albumin - creatinine ratio; WBC = white blood cell count; WC = waist circumference; WHR = waist-to-hip ratio.

Table S1: Factor analyzes of the MBS in the child and adolescent cohort

Study	Characterization	Approach	Description factors
Chen <i>et al.</i> [93]	- Bogalusa Heart Study - Biracial (black/white) community-based population - $n = 4,522$ children, adolescents, and young adults	Little jiffy	1. trig., HDL chol., FG, FI, PIn 2. SBP, DBP, FI
Chen <i>et al.</i> [94]	- Bogalusa Heart Study - Biracial (black/white) community-based population - $n = 264$ (132 whites, 132 African Americans)	Little jiffy, ethnic subgroup analyzes	<i>Whites</i> 1. %BF, SBP, DBP 2. %BF, HDL chol., trig. 3. %BF, IR, RA  <i>African Americans</i> 1. %BF, SBP, DBP 2. HDL chol., trig. 3. %BF, IR
Dwyer <i>et al.</i> [95]	- Tasmanian Infant Health Study - Children from 1989 singleton births, Tasmania - $n = 298$ 8-year-olds	EFA, factor retention based on LR testing, extraction by principal axis factoring, Harris-Kaiser rotation (oblique)	1. SBP,DBP 2. FI, FG 3. trig., HDL chol., LDL chol.
Moreno <i>et al.</i> [96]	- Schoolchildren, Zaragoza, Spain - $n = 142$ (74 non-obese, 68 obese)	Little jiffy, obesity subgroup analyzes	<i>Non-obese</i> 1. WC, BMI, FI, trig., TTTS, HDL chol., Leptin, UA 2. DBP, SBP 3. trig., HDL chol., Leptin, UA 4. FI, FG  <i>Obese</i> 1. BMI, WC, Leptin, DBP, SBP 2. DBP, TTTS, UA, SBP 3. FI, FG, trig. 4. HDL chol., trig.
Lambert <i>et al.</i> [36]	- Quebec Child and Adolescent Health and Social Survey - Quebecian youth aged 9, 13, and 16 - $n = 2,223$ (700 9-y-olds, 716 13-y-olds, 817 16-y-olds)	Little jiffy, with relaxed Guttman-Kaiser rule, age subgroup analyzes	<i>9-y-olds</i> 1. BMI, FI, FG 2. BMI, FG, trig., HDL chol. 3. SBP, DBP  <i>13-y and 16-y-olds</i> 1. BMI, FI, trig., HDL chol. 2. SBP, DBP 3. FI, FG
Park <i>et al.</i> [97]	- School-going volunteers, urban South Korea - Middle and high-school students aged 13-18 - $n = 148$ (68 boys, 80 girls)	PCA, Guttman-Kaiser rule, unspecified orthogonal rotation, gender subgroup analyzes	<i>Boys</i> 1. BMI, %BF, WC, LDL chol., trig., Leptin 2. SBP, DBP 3. FG, LDL chol., HDL chol.  <i>Girls</i> 1. BMI, %BF, WC, Leptin, FG 2. SBP, DBP 3. LDL chol., HDL chol., trig.

Table S1 (Continued)

Study	Characterization	Approach	Description factors
Weiss <i>et al.</i> [69]	<ul style="list-style-type: none"> <li>- Multiethnic, multiracial cohort</li> <li>- Nondiabetic overweight children and adolescents</li> <li>- <math>n = 464</math></li> </ul>	Little jiffy	<ol style="list-style-type: none"> <li>1. BMI, HOMA-IR, FG, PG</li> <li>2. trig., HDL chol.</li> <li>3. SBP, DBP</li> </ol>
Goodman <i>et al.</i> [98]	<ul style="list-style-type: none"> <li>- Princeton School District Study</li> <li>- School-based study, Cincinnati, Ohio, USA</li> <li>- Nondiabetic 7th to 12th graders</li> <li>- <math>n = 212</math> (subsample with BP measurements)</li> </ul>	Little jiffy	<ol style="list-style-type: none"> <li>1. FG, WC, BMI, fib.</li> <li>2. LDL chol. TC</li> <li>3. FG, FI, trig., HDL chol.</li> <li>4. SBP, DBP</li> </ol>
Retnakaran <i>et al.</i> [99]	<ul style="list-style-type: none"> <li>- Sandy Lake Health and Diabetes Project</li> <li>- Nondiabetic children aged 10-19</li> <li>- <math>n = 231</math></li> </ul>	EFA, factor retention based on scree plot, extraction by principal axis factoring, oblique rotation (Promax)	<ol style="list-style-type: none"> <li>1. WC, %BF, BMI, trig., FI</li> <li>2. trig., HDL chol., FG, PG, FI</li> <li>3. SBP, DBP</li> </ol>
Ghosh [100]	<ul style="list-style-type: none"> <li>- Calcutta, India</li> <li>- Apparently healthy Asian Indian adolescents</li> <li>- <math>n = 400</math> (200 boys, 200 girls)</li> </ul>	Little jiffy, gender subgroup analyzes	<p><i>Boys</i></p> <ol style="list-style-type: none"> <li>1. WC</li> <li>2. S:T</li> <li>3. TC, trig., FG</li> <li>4. SBP, DBP, MAP</li> </ol> <p><i>Girls</i></p> <ol style="list-style-type: none"> <li>1. WC, S:T</li> <li>2. WC, S:T</li> <li>3. TC, trig., FG</li> <li>4. SBP, DBP, MAP</li> </ol>
Goodman <i>et al.</i> [101]	<ul style="list-style-type: none"> <li>- Princeton School District Study</li> <li>- School-based study, Cincinnati, Ohio, USA</li> <li>- <math>n = 1,098</math></li> </ul>	Little jiffy, baseline and 3-year follow-up subgroup analyzes	<p><i>Baseline</i></p> <ol style="list-style-type: none"> <li>1. FI, BMI, WC</li> <li>2. FI, HDL chol., trig., FG</li> <li>3. DBP, SBP</li> </ol> <p><i>Follow-up</i></p> <ol style="list-style-type: none"> <li>1. FI, BMI, WC, FG</li> <li>2. HDL chol., trig.</li> <li>3. DBP, SBP</li> </ol>
Kelishadi <i>et al.</i> [102]	<ul style="list-style-type: none"> <li>- CASPIAN Study</li> <li>- Subjects from 6 provinces in Iran, aged 6 to 18</li> <li>- <math>n = 4,811</math> (678 ATPIII MBS, 4,133 non-MBS)</li> </ul>	Little jiffy, subgroup analyzes by age and MBS-designation	<p><i>Non-MBS</i></p> <ol style="list-style-type: none"> <li>1. TC, LDL chol.</li> <li>2. trig., HDL chol., FG</li> <li>3. WC, BMI</li> <li>4. DBP, SBP</li> </ol> <p><i>MBS</i></p> <ol style="list-style-type: none"> <li>1. TC, LDL chol., trig.</li> <li>2. HDL chol., FG, WC, BMI</li> <li>3. WC, DBP, SBP</li> </ol>
Li <i>et al.</i> [103]	<ul style="list-style-type: none"> <li>- National Health and Nutrition Examination Survey</li> <li>- Adolescents aged 12-17</li> <li>- <math>n = 1,262</math></li> </ul>	CFA, ML estimation, assessment measurement invariance across sex and	<ol style="list-style-type: none"> <li>1. WC, trig., FI, SBP</li> </ol> <p>(Model considered invariant)</p>

Table S1 (Continued)

Study	Characterization	Approach ethnicity	Description factors
Ng <i>et al.</i> [104]	<ul style="list-style-type: none"> <li>- Hong-Kong Chinese adolescents</li> <li>- Aged 12-19</li> <li>- <math>n = 2,102</math> (958 boys, 1,144 girls)</li> </ul>	Little jiffy, gender subgroup analyzes	<p><i>Boys</i></p> <ol style="list-style-type: none"> <li>1. BW, WC, BMI</li> <li>2. SBP, DBP</li> <li>3. trig., HDL chol., LDL chol.</li> <li>4. FG</li> </ol> <p><i>Girls</i></p> <ol style="list-style-type: none"> <li>1. BW, WC, BMI</li> <li>2. SBP, DBP, FG</li> <li>3. trig., HDL chol.</li> <li>4. LDL chol.</li> </ol>
Ramachandran <i>et al.</i> [105]	<ul style="list-style-type: none"> <li>- School-based survey, Chennai, India</li> <li>- Children aged 12-19</li> <li>- <math>n = 2,640</math> (1,323 boys, 1,317 girls)</li> </ul>	PCA, component retention and rotation (orthogonal) unspecified, gender subgroup analyzes	<p><i>Boys</i></p> <ol style="list-style-type: none"> <li>1. WC, SBP, DBP</li> <li>2. trig. WC, FG, FI</li> <li>3. HDL chol., trig., FG</li> </ol> <p><i>Girls</i></p> <ol style="list-style-type: none"> <li>1. WC, SBP, DBP, FI</li> <li>2. HDL chol., trig., WC, FI</li> <li>3. FG, FI</li> </ol>
Goodman <i>et al.</i> [56]	<ul style="list-style-type: none"> <li>- Fels Longitudinal Study</li> <li>- Focus on developmental puberty stages</li> <li>- <math>n = 442</math></li> </ul>	CFA, assessment invariance of 3 alternative models (1-factor model, 4-factor model, 2nd-order factor model) across developmental puberty stages	None of the models gave (consistent) adequate fit (across
Mayer-Davis <i>et al.</i> [106]	<ul style="list-style-type: none"> <li>- SEARCH for Diabetes in Youth Study</li> <li>- Type 1 and Type 2 diabetics, aged 10-22</li> <li>- <math>n = 1,293</math> (1,198 Type 1, 95 Type 2)</li> </ul>	CFA, ML estimation, assessment various CFA models (1-3 factors), assessment measurement invariance across diabetes type	<ol style="list-style-type: none"> <li>1. BMI, WC</li> <li>2. trig., HDL chol.</li> <li>3. DBP, SBP</li> </ol> (Model considered invariant)
Noto <i>et al.</i> [107]	<ul style="list-style-type: none"> <li>- Primary and secondary schools, Serre Calabre Montane, Italy</li> <li>- Schoolchildren aged 7-14</li> <li>- <math>n = 1,629</math> (859 boys, 770 girls)</li> </ul>	Little jiffy, gender subgroup analyzes	<ol style="list-style-type: none"> <li>1. WC, BMI, age, SBP</li> <li>2. HDL chol. trig.</li> <li>3. SBP, DBP</li> <li>4. FG</li> </ol> (Structure considered invariant)
Ghasemi <i>et al.</i> [108]	<ul style="list-style-type: none"> <li>- Tehran Lipid and Glucose Study</li> <li>- District 13 of Tehran, Iran</li> <li>- <math>n = 851</math> (88 with MBS, 236 overweight)</li> </ul>	Little jiffy, subgroup analyzes w.r.t. MBS-designation and weight	<p><i>Total sample</i></p> <ol style="list-style-type: none"> <li>1. SBP, DBP, WC, BMI</li> <li>2. trig., HDL chol., WC, BMI</li> <li>3. FG, NO<sub>x</sub></li> </ol> <p><i>MBS</i></p> <ol style="list-style-type: none"> <li>1. WC, BMI</li> <li>2. trig., HDL chol.</li> <li>3. SBP, DBP</li> <li>4. FG, NO<sub>x</sub></li> </ol>

Table S1 (Continued)

Study	Characterization	Approach	Description factors
			<i>Overweights</i> 1. SBP, DBP, WC 2. trig., HDL chol., WC 3. FG, BMI, NO <sub>x</sub>
Martínez-Vizcaíno <i>et al.</i> [109]	- Children from 20 schools in Cuenca, Spain - Children aged 10-13 - $n = 1,020$	CFA, usage of parceling, assessment measurement invariance across gender and physical activity	1. WC, trig./HDL chol., FI, MAP (Model considered invariant)
Khader <i>et al.</i> [43]	- Household sample from 12 governorates, Jordan - Children and adolescents - $n = 655$	Little jiffy, subgroup analyzes w.r.t. gender and age	1. BMI, WC 2. DBP, SBP 3. HDL chol., trig. 4. FG (Structure considered invariant)
Martínez-Vizcaíno <i>et al.</i> [110]	- European Youth Heart Study - 6-year follow-up study - Swedish and Estonian children aged 9 at start - $n = 634$ (174 Swedish, 460 Estonian)	CFA, ML estimation, usage of parceling, assessment measurement invariance across ethnicity and time	1. WC, trig./HDL chol., FI, MAP (Model considered invariant)
Olza <i>et al.</i> [111]	- Children from 3 regions of Spain - Obese Caucasian children aged 5-14 - $n = 478$	Little jiffy	1. SBP, FG, HOMA-IR 2. HDL chol., trig. 3. BMI, SBP
Gurka <i>et al.</i> [112]	- National Health and Nutrition Examination Survey - Non-Hispanic blacks, whites, and Hispanics aged 12-19 - $n = 4,147$	CFA, ML estimation, assessment measurement invariance across sex-ethnicity combinations	1. BMI, SBP, HDL chol., trig., FG (Structure considered invariant)
Hong <i>et al.</i> [113]	- Ho Chi Minh City, Vietnam - Urban high-school students - $n = 617$ (284 boys, 333 girls)	Little jiffy, gender subgroup analyzes	<i>Boys</i> 1. BMI, WC 2. SBP, DBP, LDL chol. 3. trig., HDL chol.  <i>Girls</i> 1. BMI, WC 2. SBP, DBP, LDL chol. 3. trig., HDL chol. 4. LDL chol., FG
Suchday <i>et al.</i> [114]	- Students liberal arts college, Mumbai, India - Students aged 15-23 - $n = 112$	Little jiffy	1. SBP, WHR, DBP, BMI, FI 2. TC, trig. 3. FG, BMI, FI
Wang <i>et al.</i> [115]	- Beijing Child and Adolescent Metabolic Syndrome Study - Schoolchildren aged 6 to 18	Little jiffy, gender subgroup analysis,	<i>Non-parceling, boys</i> 1. WC, trig., HDL chol., HOMA-IR, FI, Leptin, Adi

Table S1 (Continued)

Study	Characterization	Approach	Description factors
	- $n = 3,373$ (1,717 boys, 1,656 girls)	followed by parceling	2. WC, SBP, DBP 3. FG, HOMA-IR, FI  <i>Non-parceling, girls</i> 1. WC, FG, HOMA-IR, FI, Leptin 2. WC, trig., HDL chol., Leptin, Adi 3. WC, SBP, DBP  <i>Usage of parceling</i> 1. WC, MAP, trig./HDL chol., Leptin/Adi

Table S2: Factor analyzes of the MBS in additional cohorts

Study	Characterization	Approach	Description factors
Edwards <i>et al.</i> [26,116]	- Kaiser-Permanente Women Twins Study - Unrelated, nondiabetic women - $n = 281$	Little jiffy	1. BW, WC, FI, FG 2. FI, PI, FG, PG, SBP 3. trig., HDL chol., LDL-PPD
Donahue <i>et al.</i> [117]	- Miami Community Health Study - Nondiabetic African-Americans and non-Hispanic whites - $n = 50$	Little jiffy	1. UA, SBP, DBP, HDL chol., trig., WC, IMGD 2. FI, FG, DBP, IMGD
Meigs <i>et al.</i> [118]	- Framingham Offspring Study - Nondiabetic men and women - $n = 2,458$ (1,150 men and 1,308 women)	Little jiffy, gender subgroup analyzes	1. FI, PI, BMI, WHR, HDL chol., trig. 2. GB, G2, FI, PI 3. BMI, SBP, DBP (Structure considered invariant)
Edwards <i>et al.</i> [119]	- Honolulu Heart Program - (Non)diabetic elderly Japanese-American men - $n = 3,159$ (2,760 nondiabetic and 399 diabetic)	Little jiffy, subgroup analyzes according to diabetic status	<i>Nondiabetic</i> 1. BW, WC, FI 2. SBP, DBP 3. trig., HDL chol. 4. FG, FI  <i>Diabetic</i> 1. BW, WC 2. SBP, DBP 3. trig., HDL chol., FG 4. FG, FI
Gray <i>et al.</i> [120]	- Strong Heart Study - American Indians - $n = 4,228$ (975 and 1,202 nondiabetic men and women, 783 and 1,268 diabetic men and women)	Little jiffy, subgroup analyzes by gender and diabetic status	<i>Nondiabetic males</i> 1. BMI, FI, FG 2. SBP, DBP 3. HDL chol., trig., FI  <i>Nondiabetic females</i> 1. BMI, FI, FG 2. SBP, DBP 3. HDL chol., trig.  <i>Diabetic males</i> 1. SBP, DBP 2. HDL chol., trig., FG



Table S2 (Continued)

Study	Characterization	Approach	Description factors
			3. HDL chol., BMI, FI
			<i>Diabetic females</i> 1. SBP, DBP 2. HDL chol., trig., FG 3. BMI, FI
Leyva <i>et al.</i> [121]	- RISC-2 Study - Nondiabetic men - $n = 74$	Little jiffy	1. IR, FG, FI, Leptin 2. Leptin, IVGTT, S:T, UA, SBP, BMI 3. HDL chol., trig.
Chen <i>et al.</i> [122]	- Nondiabetic residents Kinmen Taiwan - $n = 8,437$ (3,659 men and 4,778 women)	Little jiffy, gender subgroup analyzes	<i>Males</i> 1. SBP, DBP 2. FI, BMI, WHR, HDL chol., trig. 3. FG  <i>Females</i> 1. SBP, DBP 2. FI, BMI, WHR, HDL chol., trig. 3. FG, trig.
Sakkinen <i>et al.</i> [123]	- Cardiovascular Health Study - Elderly nondiabetics - $n = 322$	Little jiffy	1. BW, WC, FI, FG 2. PI, PG 3. SBP, DBP 4. trig., HDL chol.
Snehalatha <i>et al.</i> [124]	- Urban cluster survey, Madras, India - Nondiabetic adults - $n = 654$ (396 men and 258 women)	PCA, component retention based on scree plot, Varimax rotation, gender subgroup analyzes	<i>Males</i> 1. BMI, FG, PI, HOMA-IR 2. BMI, SBP, DBP 3. BMI, WHR, chol., trig.  <i>Females</i> 1. BMI, FG, PI, HOMA-IR 2. BMI, SBP, DBP 3. chol., trig. 4. BMI, WHR, HOMA-IR
Anderson <i>et al.</i> [125]	- Chinese men and women - $n = 145$	EFA, factor retention based on Guttman-Kaiser rule, Varimax rotation	1. FI, FG, BMI, WC, IS 2. BMI, MAP, WC 3. HDL chol., WC, trig.
Hodge <i>et al.</i> [126]	- Island of Mauritius - Nondiabetic adults - $n = 3,068$ (1,414 men and 1,654 women)	Little jiffy, gender subgroup analyzes	<i>Males</i> 1. WHR, BMI, Leptin, trig., HDL chol., FA, PI 2. UA, SBP, DBP 3. FI, PI, FG, PG  <i>Females</i> 1. WHR, BMI, Leptin, trig., HDL chol., FI, PI, UA 2. FG, SBP, DBP 3. FI, PI, FG, PG

Table S2 (Continued)

Study	Characterization	Approach	Description factors
Shmulewitz <i>et al.</i> [127]	- Pacific island of Kosrae, Micronesia - Nondiabetic adults - $n = 628$	PCA, component retention and rotation (orthogonal) unspecified	1. BW, WHR, Leptin, FI 2. TC, trig., Apo B 3. FG, SBP, DBP 4. Leptin, trig., Apo A-I
Adami <i>et al.</i> [128]	- Obese Italian adults - $n = 163$	Little jiffy	1. trig., HDL chol., SBP, DBP 2. FI, HOMA-IR 3. BMI, Leptin, FG
Arya <i>et al.</i> [129]	- San Antonio Family Diabetes Study - Nondiabetic Mexican-Americans - $n = 261$	Little jiffy	1. BMI, Leptin, FI 2. DBP, SBP 3. HDL chol., trig.
Hanley <i>et al.</i> [33]	- Insulin Resistance Atherosclerosis Study - Nondiabetic adults - $n = 1,087$	Little jiffy	1. BMI, WC, FG, PG, HOMA-IR, trig., HDL chol. 2. DBP, SBP (See [33] for subgroup analyzes)
Hanson <i>et al.</i> [34]	- Gila River Indian Community, Arizona, USA - Adult Pima Indians - $n = 1,448$ (1,918 nondiabetics, 470 diabetics)	Little jiffy, usage of parceling, subgroup analyzes by diabetic status	<i>Nondiabetics</i> 1. FI×FG, PI×PG, FI/FG, PI/PG 2. BW, WC 3. DBP, SBP 4. HDL chol., trig.  <i>Diabetics</i> 1. FI×FG, PI×PG, FI/FG, PI/PG 2. BW, WC 3. DBP, SBP 4. HDL chol., trig., FI×FG
Kue Young <i>et al.</i> [130]	- Sample survey 3 contiguous regions of Canada - Adult Indians, Inuit, and non-Aboriginal Canadians - $n = 3,930$ (873 Indians, 387 Inuit, 2,670 non-Aborigines)	Little jiffy	1. BMI, HC, WC 2. SBP, DBP, TC 3. trig., HDL chol., FG
O'Dea <i>et al.</i> [131]	- Aboriginal and Torres Strait Islander People, Australia - Community-based risk factor survey - $n = 863$ (643 Aborigines and 220 Islanders)	Little jiffy, subgroup analyzes by gender and ethnicity	<i>TSI, males</i> 1. FG, PG, FI, trig., HDL chol., WHR 2. FI, SBP, DBP, BMI, WHR  <i>TSI, females</i> 1. FG, PG, FI, trig., HDL chol., BMI, WHR 2. SBP, DBP  <i>Aborigines, males</i> 1. SBP, DBP, BMI, WHR 2. FG, PG, BMI, WHR 3. FI, trig., HDL chol., BMI  <i>Aborigines, females</i> 1. SBP, DBP 2. FG, PG 3. FI, trig., HDL chol., BMI, WHR

Table S2 (Continued)

Study	Characterization	Approach	Description factors
Choi <i>et al.</i> [132]	<ul style="list-style-type: none"> <li>- The South-West Seoul Study</li> <li>- Urban elderly nondiabetic Koreans</li> <li>- <math>n = 1,314</math> (249 men, 1,065 women)</li> </ul>	Little jiffy, gender subgroup analyzes	<p><i>Males</i></p> <ol style="list-style-type: none"> <li>1. BMI, WHR, HDL chol., trig., PI, FI</li> <li>2. SBP, DBP</li> <li>3. PI, PG</li> <li>4. FI, FG</li> </ol> <p><i>Females</i></p> <ol style="list-style-type: none"> <li>1. SBP, DBP</li> <li>2. BMI, WHR, HDL chol., trig.</li> <li>3. FG, PG, PI</li> <li>4. BMI, WHR, FI, FG</li> </ol>
Ford [35]	<ul style="list-style-type: none"> <li>- National Health and Nutrition Examination Survey</li> <li>- <math>n = 6,868</math> (3,410 men, 3,458 women)</li> </ul>	Little jiffy, subgroup analyzes by gender, ethnicity and age	<p><i>Males</i></p> <ol style="list-style-type: none"> <li>1. HDL chol., trig., FI, WC, BMI, WHR, UA</li> <li>2. SBP, DBP</li> <li>3. FG, U:C</li> </ol> <p><i>Females</i></p> <ol style="list-style-type: none"> <li>1. FI, WC, BMI, WHR, UA</li> <li>2. SBP, DBP</li> <li>3. HDL chol., trig., FG</li> </ol> <p>(See [35] for other subgroup analyzes)</p>
Hanley <i>et al.</i> [133]	<ul style="list-style-type: none"> <li>- Sandy Lake Health and Diabetes Project</li> <li>- Nondiabetic adult native Canadians</li> <li>- <math>n = 728</math> (305 men, 423 women)</li> </ul>	EFA, retention based on (ao) scree plot, factor extraction by principal axis factoring, Varimax rotation	<ol style="list-style-type: none"> <li>1. FI, WC, %BF</li> <li>2. FI, WC, trig., HDL chol., Adi</li> <li>3. FI, FG, PG</li> </ol>
Howard <i>et al.</i> [134]	<ul style="list-style-type: none"> <li>- Women's Health Initiative</li> <li>- Nondiabetic postmenopausal female adults</li> <li>- <math>n = 3,083</math> (1,635 White, 802 Black, 390 Latino, 256 Asian/Pacific Islander)</li> </ul>	Little jiffy, ethnicity subgroup analyzes	<p><i>Whites</i></p> <ol style="list-style-type: none"> <li>1. BMI, HC, WC, HOMA-IR, FI, FG</li> <li>2. HOMA-IR, FI, HDL chol., trig., HDL<sub>2</sub>/HDL tot.</li> <li>3. TC, LDL chol.</li> <li>4. SBP, DBP</li> </ol> <p>(Structure similar in other ethnic subgroups)</p>
Novak <i>et al.</i> [135]	<ul style="list-style-type: none"> <li>- Göteborg, Sweden</li> <li>- Middle-aged Caucasian males</li> <li>- <math>n = 284</math></li> </ul>	CFA, ML estimation	<ol style="list-style-type: none"> <li>1. SBP, DBP</li> <li>2. BMI, WHR</li> <li>3. FI, FG</li> <li>4. trig., HDL chol.</li> </ol>
Shen <i>et al.</i> [52]	<ul style="list-style-type: none"> <li>- Normative Aging Study</li> <li>- Male participants</li> <li>- <math>n = 847</math></li> </ul>	CFA, robust ML estimation, second-order factor taken to reflect MBS, assessment measurement invariance over cardiovascular disease and age groupings	<ol style="list-style-type: none"> <li>1. FI, PI, FG, PG</li> <li>2. BMI, WHR</li> <li>3. trig., HDL chol.</li> <li>4. SBP, DBP</li> </ol> <p>(Structure deemed invariant)</p>
Tang <i>et al.</i> [136]	<ul style="list-style-type: none"> <li>- NHLBI Family Heart Study</li> <li>- Probands' family members</li> <li>- <math>n = 2,831</math></li> </ul>	EFA, factor retention based on on Guttman-Kaiser rule, factor extraction by ML estimation, rotation unspecified	<ol style="list-style-type: none"> <li>1. BMI, WHR, SSK, trig., HDL chol., HOMA-IR, PAI-1, UA</li> <li>2. trig., LDL chol., TC</li> <li>3. BMI, SSK</li> <li>4. trig., LDL chol.</li> </ol>

Table S2 (Continued)

Study	Characterization	Approach	Description factors
Cai <i>et al.</i> [137]	- San Antonio Family Heart Study - Nondiabetic Mexican-Americans - $n = 566$	Little jiffy	1. BW, BMI, WC, HC, %BF, FI, AST 2. FG, PG, FI, PI 3. SBP, DBP 4. HDL chol., trig.
Jones <i>et al.</i> [138]	- Spinal cord-injured men - $n = 20$	Little jiffy	1. %BF, PAL, TFM, PG, PI 2. PAL, TC/HDL chol., HDL chol., PI 3. FG, FI
Nelson <i>et al.</i> [139]	- Swedish Adoption/Twin Study of Aging - $n = 1,944$	CFA, further approach unspecified, usage of parceling	1. IS, MAP, WHR, TC/HDL chol.
Oh <i>et al.</i> [37]	- Mokdong Study of Diabetes Prevalence - Urban Korean adults - $n = 655$ (206 men, 449 women)	PCA, component retention unspecified, Varimax rotation, gender subgroup analyzes	<i>Males</i> 1. BMI, WC, FI 2. FG, PG, FI 3. SBP, DBP 4. FI, trig., HDL chol.  <i>Females</i> 1. BMI, WC, SBP, DBP 2. FG, PG, FI 3. BMI, WC, FI, trig., HDL chol.
Wang <i>et al.</i> [140]	- National Diabetes Survey China - Nondiabetic Chinese adults - $n = 934$ (449 and 485 nondiabetic men and women, 132 and 173 diabetic men and women)	Little jiffy, gender and diabetic status subgroup analyzes	<i>Nondiabetic males</i> 1. BMI, WHR, FI, FG 2. SBP, DBP 3. PI, FG, PG 4. TC, trig.  <i>Nondiabetic females</i> 1. BMI, WHR, FI 2. SBP, DBP 3. PI, FG, PG 4. TC, trig.  <i>Diabetic males</i> 1. SBP, DBP 2. UA, trig., FG, PG 3. BMI, WHR, FI 4. TC, FI, PI  <i>Diabetic females</i> 1. SBP, DBP 2. UA, FG, PG 3. BMI, WHR, trig., FI 4. TC, FI, PI
Ang <i>et al.</i> [38]	- Singapore National Health Survey - Nondiabetic Chinese, Malay, and Asian Indian adults - $n = 4,265$ (1,671 Chinese, 281 Malay, 158 Indian)	Little jiffy, subgroup analyzes w.r.t. gender and ethnicity	<i>Chinese males and females, Malay males</i> 1. HOMA-IR, BMI, WHR, HDL chol., trig. 2. FG, PG, HOMA-IR

Table S2 (Continued)

Study	Characterization	Approach	Description factors
	males; 1706 Chinese, 264 Malay, 146 Indian females)		3. BMI, WHR, SBP, DBP
			<i>Malay females</i> 1. HOMA-IR, BMI, WHR, HDL chol., trig. 2. FG, PG, HOMA-IR 3. HOMA-IR, BMI, WHR, trig., SBP, DBP
			<i>Indian males</i> 1. HOMA-IR, BMI, WHR, HDL chol., trig. 2. FG, PG, HOMA-IR 3. WHR, SBP, DBP
			<i>Indian females</i> 1. HOMA-IR, BMI, WHR, HDL chol., trig. 2. FG, PG, HOMA-IR, WHR 3. BMI, WHR, SBP, DBP
Ghosh [39]	- Employees Eastern Railway Government, India - Middle-aged Bengalee Hindu men - $n = 212$	Little jiffy	1. WHR 2. TER 3. TC, trig., FG 4. SBP, DBP
Lin <i>et al.</i> [141]	- Northern Manhattan Family Study - Subjects from Caribbean-Hispanic families - $n = 803$	Little jiffy	1. trig., HDL chol., FG, WC 2. SBP, DBP
Mohan <i>et al.</i> [142]	- Chennai Urban Rural Epidemiology Study - Adult diabetic from Chennai, India - $n = 100$	Little jiffy, component retention not specified	1. trig., HDL chol., HOMA-IR, Adi, WC 2. WC, SBP, DBP
Liou <i>et al.</i> [143]	- Taipei Veterans General Hospital - Cross-sectional studies from Spain, Mauritius, and USA - $n = 393$	Little jiffy	1. BMI, WC, FG, trig., HDL chol., UA 2. SBP, DBP
Pladevall <i>et al.</i> [53]	- CFA study using previously collected data - Nondiabetic middle-aged Chinese men - $n = 4,318$	CFA, ML estimation, usage of parceling	1. WC, HOMA-IR, trig./HDL chol., MAP
Shah <i>et al.</i> [54]	- Insulin Resistance Atherosclerosis Study - Data as used in [33]	CFA, ML estimation	1. SBP, DBP 2. HDL chol., trig. 3. IS, PG, FG 4. WC, BMI
Shen <i>et al.</i> [144]	- Miami Community Health Study - Nondiabetic Caucasians, African and Cuban-Americans - $n = 517$	CFA, ML estimation, second-order factor taken to reflect MBS, assessment measurement invariance over gender, ethnicity and age groupings	1. FI, FG 2. BMI, WC 3. trig., HDL chol. 4. SBP, DBP (Structure deemed invariant)

Table S2 (Continued)

Study	Characterization	Approach	Description factors
Zanolin <i>et al.</i> [145]	<ul style="list-style-type: none"> <li>- Outpatients, Verona, Italy</li> <li>- Caucasian nondiabetic hyperandrogenic women</li> <li>- <math>n = 255</math></li> </ul>	Little jiffy	<ol style="list-style-type: none"> <li>1. FI, PI, BMI, HDL chol., trig., UA</li> <li>2. BMI, DBP, SBP, FT</li> <li>3. FG, PG, PI, trig., FT</li> <li>4. PI, FT, 17HP</li> </ol>
Chien <i>et al.</i> [146]	<ul style="list-style-type: none"> <li>- Chin-Shan Community Family Study</li> <li>- Adolescent probands and relatives, China</li> <li>- <math>n = 1,227</math></li> </ul>	Little jiffy	<ol style="list-style-type: none"> <li>1. SBP, DBP, WC, BMI</li> <li>2. TC, LDL chol.</li> <li>3. HOMA-IR, FG, HDL chol., trig.</li> </ol>
Razak <i>et al.</i> [147]	<ul style="list-style-type: none"> <li>- Study of Health Assessment and Risk in Ethnic Groups</li> <li>- Subjects from 4 ethnic groups from 4 regions, Canada</li> <li>- <math>n = 1,078</math></li> </ul>	PCA, retention and rotation criteria unspecified	<ol style="list-style-type: none"> <li>1. FG, PG, FI, PI, HOMA-IR, HbA1c, FFFA, PFFA</li> <li>2. FI, PI, HOMA-IR, LDL chol., HDL chol., trig., PFFA</li> <li>3. SBP, DBP</li> </ol>
Ghosh [148]	<ul style="list-style-type: none"> <li>- Nondiabetic pre and postmenopausal Bengalee women</li> <li>- <math>n = 200</math> (100 pre and 100 postmenopausal)</li> </ul>	Little jiffy, menopausal subgroup analyzes	<p><i>Premenopausal</i></p> <ol style="list-style-type: none"> <li>1. WC, WHR</li> <li>2. TC, trig., HDL chol., LDL chol., FG</li> <li>3. SBP, DBP</li> </ol> <p><i>Postmenopausal</i></p> <ol style="list-style-type: none"> <li>1. WC, WHR</li> <li>2. WC, WHR, FG</li> <li>3. TC, trig., HDL chol., LDL chol., FG</li> <li>4. SBP, DBP</li> </ol>
Lafortuna <i>et al.</i> [149]	<ul style="list-style-type: none"> <li>- Patients IAI, Piancavallo, Italy</li> <li>- Adult obese women</li> <li>- <math>n = 552</math></li> </ul>	Little jiffy	<ol style="list-style-type: none"> <li>1. FI, HOMA-IR</li> <li>2. WHR, trig., HDL chol., FG</li> <li>3. BW, WC</li> <li>4. SBP, DBP</li> </ol>
Reimann <i>et al.</i> [150]	<ul style="list-style-type: none"> <li>- Transition and Health during Urbanization Study</li> <li>- Rural, semiurban, and urban Black South Africans</li> <li>- <math>n = 448</math> (140 rural, 118 semiurban, 190 urban)</li> </ul>	Little jiffy, subgroup analyzes w.r.t. urbanization	<p><i>Rural</i></p> <ol style="list-style-type: none"> <li>1. BMI, WC, FI, HDL chol.</li> <li>2. SBP, DBP, FI</li> <li>3. trig., UA</li> <li>4. FG, PG, HDL chol.</li> </ol> <p><i>Semiurban</i></p> <ol style="list-style-type: none"> <li>1. BMI, WC, trig.</li> <li>2. SBP, DBP, HDL chol.</li> <li>3. FG, FI</li> <li>4. PG, HDL chol.</li> <li>5. trig., UA</li> </ol> <p><i>Urban</i></p> <ol style="list-style-type: none"> <li>1. BMI, WC, HDL chol.</li> <li>2. FG, FI, trig.</li> <li>3. SBP, DBP</li> <li>4. FG, PG, HDL chol.</li> </ol>
Wu <i>et al.</i> [151]	<ul style="list-style-type: none"> <li>- Patients enrolled in a General Hospital, Taipei, Taiwan</li> <li>- Participants in routine health check</li> </ul>	Little jiffy, glucose tolerance subgroup	<p><i>Normal glucose tolerance</i></p> <ol style="list-style-type: none"> <li>1. SBP, DBP</li> </ol>

Table S2 (Continued)

Study	Characterization	Approach	Description factors
	- $n = 509$ (345 normal and 164 impaired glucose tolerance)	analyzes	2. WHR, SSPG, trig., HDL chol.
			<i>Impaired glucose tolerance</i> 1. SBP, DBP 2. WHR, SSPG, trig., HDL chol. 3. FG, trig.
Boronat <i>et al.</i> [57]	- Telde Study - Nondiabetic adult Canadians - $n = 902$	CFA, ML estimation, usage of parceling, assessment measurement invariance across gender	1. WC, trig./HDL chol., HOMA-IR, MAP (Structure deemed invariant)
Leone <i>et al.</i> [152]	- Paris Investigations Préventives et Cliniques Center - Subjects underwent health examination - $n = 121,965$	PCA, retention based on scree plot, Varimax rotation	1. HDL chol., trig. 2. SBP, FG 3. WC
Öhrvik <i>et al.</i> [153]	- 75 year olds from Västerås, Sweden - $n = 401$ (198 men, 203 women)	EFA, factor extraction by principal axis factoring, factor retention based on scree plot, Varimax rotation, gender subgroup analyzes	1. WC, HDL chol., trig., FG 2. SBP, DBP (Structure deemed invariant)
Chang <i>et al.</i> [154]	- Nondiabetic dioxin-exposed Taiwanese - $n = 1,409$	Little jiffy	1. HDL chol., trig., TC/HDL chol. 2. SBP, DBP, serum PCDD/Fs 3. BW, WC 4. FG, HOMA-IR
Esteghamati <i>et al.</i> [155]	- SuRFNCD 2007, Iran - Iranian adults - $n = 3,001$ (1,483 men, 1,518 women)	Little jiffy, gender subgroup analyzes	<i>Males</i> 1. WC, HOMA-IR, SBP, Leptin 2. trig., HDL chol.  <i>Females</i> 1. WC, SBP, Leptin 2. trig., HDL chol.
Marsland <i>et al.</i> [156,157]	- University of Pittsburgh Adult Health and behavior Project - Non-hispanic White and African American subjects - $n = 645$ middle-aged community volunteers	CFA, robust ML estimation, second-order factor taken to reflect MBS	1. SBP, DBP 2. FI, FG 3. BMI, WC 4. HDL chol., trig.
Barbosa-Leiker <i>et al.</i> [158]	- Spokane Heart Study - $n = 604$	CFA, robust ML estimation, assessment measurement invariance over time points	1. BMI, DBP, FG, HDL chol., trig. (Model considered stable across time)
Meshkani <i>et al.</i> [159]	- Unrelated Iranian adults - Normal and impaired glucose tolerant subjects - $n = 501$ (266 men, 235 women)	Little jiffy, gender subgroup analyzes	<i>Males</i> 1. BMI, WC, FI, UA 2. SBP, DBP 3. FG, trig., HDL chol.

Table S2 (Continued)

Study	Characterization	Approach	Description factors
			<i>Females</i> 1. SBP, DBP 2. BMI, WC, FI 3. FG, trig., HDL chol., UA
Solera-Martínez <i>et al.</i> [58]	- First-year university students - Universidad de Castilla-La Mancha, Spain - $n = 683$	CFA, ML estimation, usage of parceling, assessment measurement invariance over gender	1. HOMA-IR, trig./HDL chol., WC, MAP (Model considered invariant)
Azimi-Nezhad <i>et al.</i> [160]	- Stanislas cohort (France) and Kharasan province (Iran) - France: $n = 1,386$ (678 men, 708 women) - Iran: $n = 1,194$ (589 men, 605 women)	Extraction method unspecified, retention based on scree plot, usage Varimax rotation, subgroup analyzes w.r.t. gender and ethnicity	<i>Iranian males and females, French males</i> 1. SBP, DBP, WC 2. trig., WC, FG, HDL chol.  <i>French females</i> 1. SBP, DBP, WC, FG 2. trig., WC, HDL chol.
Esteghamati <i>et al.</i> [161]	- Outpatient clinic, Vali-Asr hospital, Iran - Iranian adults - $n = 894$ (327 nondiabetics, 567 diabetics)	Little jiffy, subgroup analyzes according to diabetic status	<i>Nondiabetics</i> 1. WC, HOMA-IR, trig., SBP, Apo B 2. HDL chol., Apo A-I  <i>Diabetics</i> 1. WC, HOMA-IR, SBP 2. HDL chol., Apo A-I 3. trig., Apo B
Stevenson <i>et al.</i> [162]	- Spokane Heart Study - Non-clinical healthy volunteers - $n = 434$	CFA, robust ML estimation	1. BMI, FG, HDL chol., trig., SBP
Chirinos <i>et al.</i> [163]	- Peruvian Study of Cardiovascular Disease Prevention - Andean Hispanic adults - $n = 2,513$	CFA, ML estimation, assessment measurement invariance over gender	1. WC, SBP, DBP, trig., FG (Structure considered invariant)
Dusseault-Belanger <i>et al.</i> [164]	- Centre Hospitalier Universitaire de Sherbrooke, Canada - Cohort with diverse medical histories - $n = 7,213$	PCA, component retention based on parallel analysis [165], rotation unspecified	1. BW, MAP, FG, HDL chol., trig. 2. MAP, FG
Gómez-Marcos <i>et al.</i> [166]	- EVIDENT study - Spanish adults - $n = 636$ (258 men, 378 women)	CFA, ML estimation, usage of parceling, assessment measurement invariance across gender	<i>Males</i> 1. BMI, trig./HDL chol., HOMA-IR, MAP  <i>Females</i> 1. WC, trig./HDL chol., HOMA-IR, MAP
Gurka <i>et al.</i> [167]	- National Health and Nutrition Examination Survey - Nondiabetic whites, blacks, and Hispanics - $n = 6,870$	CFA, ML estimation, assessment measurement invariance across gender and ethnicity	1. WC, SBP, HDL chol., trig., FG (Structure considered invariant)



Table S2 (Continued)

Study	Characterization	Approach	Description factors
Huo <i>et al.</i> [168]	- Residents Beijing, China - $n = 7,472$	CFA, ML estimation, usage of parceling, assessment measurement invariance across gender and age groups	<i>Middle-aged men</i> 1. WC, trig./HDL chol., FG, MAP  <i>Women / Young and senior men</i> 1. WC, trig., FG, SBP
Nilsson <i>et al.</i> [169]	- 75 year olds from Västerås, Sweden - $n = 396$ (196 men, 200 women)	PCA, retention based on, Guttman-Kaiser rule, variability eigenvalues assessed by bootstrapping, gender subgroup analyzes	<i>Males</i> 1. FG, HDL chol., trig., WC 2. DBP, SBP 3. FG, WBC  <i>Females</i> 1. FG, HDL chol., trig., WC, WBC 2. DBP, SBP
Sabanayagam <i>et al.</i> [170]	- Asian ethnic groups in Singapore - Middle-aged to elderly subjects - $n = 9,477$ (3,167 Chinese, 3,082 Malays, 3,228 Indians)	Little jiffy	<i>Chinese</i> 1. NFG, HbA1c 2. trig., HDL chol., BMI 3. SBP, DBP  <i>Malay</i> 1. NFG, HbA1c 2. SBP, DBP 3. trig., HDL chol.  <i>Indian</i> 1. NFG, HbA1c 2. SBP, DBP 3. trig., HDL chol. 4. BMI
Smits <i>et al.</i> [171]	- Subjects recruited by public advertisement - Nondiabetic individuals from Greater Seattle community - $n = 134$	CFA, ML estimation	1. trig., HDL chol., IS, SBP, DBP, IAF, PAI-1